University of Alberta

PRECODING FOR SPATIAL LAYER SEPARATION ON MULTI-USER MIMO DOWNLINK

by

Jia Liu

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Department of Electrical and Computer Engineering

Edmonton, Alberta

Fall 2007

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-33018-0
Our file Notre référence
ISBN: 978-0-494-33018-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

To Peng, Alexander Si You, Brendan Si Yuan, and my parents

Abstract

Multi-user wireless systems employing multiple input multiple output (MIMO) antenna techniques have recently attracted considerable research interest. A major problem in the multi-user MIMO systems is the spatial layer separation on the downlink. This thesis contributes to the design of the nonlinear pre-processing and nonlinear joint transmitter-receiver (Tx-Rx) processing algorithms for spatial layer separation on the multi-user MIMO downlink.

A generalized zero-forcing (ZF) nonlinear pre-processing structure achieving minimum bit error rate (BER) for each mobile under any given power allocation to different data streams is proposed. Based on this structure, different ZF-based nonlinear pre-processing algorithms satisfying various performance requirements are realized. Then, a minimum mean-square error (MMSE) criterion based nonlinear pre-processing algorithm is introduced. The MMSE-based algorithm mitigates the noise enhancement, and outperforms the ZF-based counterpart.

An optimal ordering lemma is introduced. This lemma gives the conditions under which the "best-first" ordering is optimal. Subsequently, it is used to solve the ordering problems in nonlinear pre-processing and nonlinear joint Tx-Rx process-

ing algorithms proposed in this thesis.

The remainder of the thesis focuses on the nonlinear joint Tx-Rx processing algorithms for the multi-user MIMO downlink with multiple-antenna mobiles. A nonlinear multi-user MIMO decomposition (NL-DECOM) technique is proposed. The application of nonlinear pre-processing at the base station pre-eliminates multi-user interference (MUI) among different mobiles. As a result, an equivalent single-user MIMO channel is created for each mobile, and self-interference among the multiple data streams directed to each individual user can be eliminated using spatial layer separation algorithms available for single-user MIMO systems. This NL-DECOM technique can achieve significantly better performance than the linear multi-user MIMO decomposition (L-DECOM) technique known in the literature. To further improve the performance, the ZF and MMSE criteria based nonlinear joint Tx-Rx processing algorithms are proposed. In these algorithms, nonlinear pre-processing is applied at the transmitter to pre-eliminate the MUI, while the self-interference is suppressed via joint nonlinear pre-processing and linear receiver processing. This approach improves performance by utilizing processing capabilities of the mobiles, while at the same time maintaining their relatively low complexity. Significant performance gain can be achieved with the application of the proposed novel algorithms.

Acknowledgements

First, I would like to express my sincere gratitude to my Ph.D. program supervisor, Dr. Witold A. Krzymień, for his invaluable guidance, fruitful discussion, enlightening suggestions, and financial support throughout the course of my doctoral research. Without him, this thesis would never have been possible.

I would also like to thank my examiners for their interest in my work and their willingness to serve on my committee.

My special gratitude goes to my husband, Dr. Peng Tan for his continuous encouragement and numerous technical discussions with me. I feel particularly indebted to my parents for their understanding, unconditional support, encouragement, and endless love throughout years.

It has been my pleasure to work at the TRLabs where I have been offered not only financial assistance but also great opportunities to conduct my research and to learn from other wonderful people.

Table of Contents

CHAPTER 1 INTRODUCTION **1**

1.1 Brief Introduction to MIMO Systems 1

1.2 Brief Introduction to Multi-User MIMO Systems 5

1.3 Precoding for Spatial Layer Separation on Multi-user MIMO Downlink and Summary of Contributions 7

1.4 Thesis Organization 12

1.5 Notation 14

CHAPTER 2 NONLINEAR PRE-PROCESSING ALGORITHMS FOR MULTI-USER MIMO DOWNLINK **16**

2.1 System Model and Assumptions 18

2.2 Introduction to the Tomlinson-Harashima Precoder 20

2.3 Conventional ZF Nonlinear Pre-Processing Algorithm 23

2.4 Generalized ZF Nonlinear Pre-Processing Structure 28

2.4.1 The Generalized Structure 29

2.4.2 Determining the Power Allocation Factors for the Generalized ZF Nonlinear Pre-Pocessing Sructure 31

2.4.3 Relationship between the Generalized ZF Nonlinear Pre-Processing Structure and the Conventional ZF Nonlinear Pre-Processing Algorithm 33

2.4.4 Simulation Results 35

2.5 MMSE Nonlinear Pre-Processing Algorithm 39

2.5.1 Comparison of the Proposed MMSE Nonlinear Pre-Processing Algorithm and Other MMSE-Criterion-Based Design 42

2.5.2 Simulation Results 44

CHAPTER 3 OPTIMAL ORDERING LEMMA 52

3.1 The Importance of Ordering 52

3.2 Change of the System Structure Due to Ordering 54

3.3 Optimal Ordering Lemma 56

3.4 Ordering in the Conventional ZF Nonlinear Pre-Processing Algorithm and the MMSE Nonlinear Pre-Processing Algorithm 59

3.4.1 An Efficient Way to Perform the "Best-First" Ordering for the Conventional ZF Nonlinear Pre-Processing Algorithm and the MMSE Nonlinear Pre-Processing Algorithm 60

3.4.2 Simulation Results 62

3.5 Ordering in Other ZF-Based Nonlinear Pre-Processing Algorithms 64

3.5.1 Simulation Results 67

CHAPTER 4 NONLINEAR JOINT TRANSMITTER-RECEIVER PROCESSING ALGORITHMS 73

4.1 System Model and Assumptions 76

4.2 Multi-User MIMO Decomposition Techniques 77

4.2.1 Linear Multi-User MIMO Decomposition Technique 77

4.2.2 Nonlinear Multi-User MIMO Decomposition Technique 78

4.2.3 Advantages of the NL-DECOM Technique over the L-DECOM Technique 82

4.2.4 Ordering in the NL-DECOM Technique 87

4.2.5 Self-Interference Cancellation Algorithms for the NL-DECOM Technique 88

4.2.6 Simulation Results 89

4.3 ZF/MMSE Nonlinear Joint Tx-Rx Processing Algorithm 96

4.3.1 Structure of the ZF/MMSE Nonlinear Joint Tx-Rx Processing Algorithm 96

4.3.2 ZF Nonlinear Joint Tx-Rx Processing Algorithm 99

4.3.3 MMSE Nonlinear Joint Tx-Rx Processing Algorithm 119

CHAPTER 5 CONCLUSIONS 131

5.1 Conclusions 131
5.2 Suggestions for Future Work 134

REFERENCES 137

APPENDIX A OPTIMAL ZF NONLINEAR PRE-PROCESSING UNDER THE RELATIVE SNR REQUIREMENT 146

APPENDIX B MMSE NONLINEAR PRE-PROCESSING ALGORITHM 147

APPENDIX C DERIVATION OF THE ERROR COVARIANCE MATRIX FOR THE MMSE NONLINEAR PRE-PROCESSING ALGORITHM 150

APPENDIX D PROOF OF THE OPTIMAL ORDERING LEMMA... 151

APPENDIX E REDUCED COMPLEXITY METHOD USED IN THE ORDERING OF THE NONLINEAR PRE-PROCESSING ALGORITHMS. 153

APPENDIX F PROOF THAT THE ZF NONLINEAR JOINT TX-RX PROCESSING ALGORITHM ACHIEVES THE MULTI-USER MIMO DECOMPOSITION 156

APPENDIX G ZF NONLINEAR JOINT TX-RX PROCESSING ALGORITHM 158

APPENDIX H OPTIMAL ORDER IN THE MAXIMAL DIVERSITY ORDER SENSE FOR THE ZF NONLINEAR JOINT TX-RX PROCESSING ALGORITHM 160

APPENDIX I APPLICATION OF THE "BEST-FIRST" ORDERING FOR THE ZF NONLINEAR JOINT TX-RX PROCESSING ALGORITHM 162

APPENDIX J DESIGN OF THE FEEDBACK FILTER FOR THE MMSE NONLINEAR JOINT TX-RX PROCESSING ALGORITHM 165

APPENDIX K DESIGN OF THE FEEDFORWARD FILTER AND DERIVATION OF THE NOISE VARIANCE FOR THE MMSE NONLINEAR JOINT TX-RX PROCESSING ALGORITHM 167

APPENDIX L PROOF THAT THE TRANSMITTED POWER IS ALLOCATED EQUALLY IN THE MMSE NONLINEAR JOINT TX-RX PROCESSING ALGORITHM..... 169

List of Figures

Figure 1.1 Basic structure of an $M \times N$ MIMO system. 3
Figure 1.2 Basic structure of a multi-user MIMO system. 6
Figure 2.1 Structure of the TH precoder for SISO channel with ISI. 20
Figure 2.2 Structure of the nonlinear pre-processing on the downlink of a multi-user MIMO system. 24
Figure 2.3 Structure of the generalized ZF nonlinear pre-processing. 31
Figure 2.4 Performance comparison of the balanced ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (BPSK, $M = N = 4$). 37
Figure 2.5 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (BPSK, $M = N = 4$). 38
Figure 2.6 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (BPSK, $M = N = 8$). 39
Figure 2.7 Performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $M = N = 4$). 47
Figure 2.8 Performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (16QAM, $M = N = 4$). 48
Figure 2.9 Performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $M = 4$, and $N = 6$). 49

Figure 2.10 Performance comparison of the proposed MMSE nonlinear pre-processing algorithm and the JBU MMSE nonlinear pre-processing algorithm (16QAM, $M = N = 4$). 51

Figure 3.1 Nonlinear pre-processing structure for the downlink of a multi-user MIMO system without ordering. 55

Figure 3.2 Nonlinear pre-processing structure for the downlink of a multi-user MIMO system if order L is used. 56

Figure 3.3 Performance comparison of the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is applied (QPSK, $M = N = 4$). 63

Figure 3.4 Performance comparison of the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is applied (16QAM, $M = N = 4$). 64

Figure 3.5 Performance comparison of the balanced ZF nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is used (BPSK, $M = N = 4$) 70

Figure 3.6 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is used (BPSK, $M = N = 4$). 71

Figure 3.7 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is used (BPSK, $M = N = 8$) 72

Figure 4.1 Structure of the NL-DECOM technique in the downlink of a multi-user MIMO system. 79

Figure 4.2 The equivalent single-user MIMO channels formed using the NL-DECOM technique 82

Figure 4.3 Performance comparison of the NL-DECOM and the L-DECOM techniques when the ZF LP or the precoded ZF DFD is used to cancel self-interference (16QAM, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$) 91

Figure 4.4 Performance comparison of the NL-DECOM and the L-DECOM techniques when the MMSE linear joint Tx-Rx processing algorithm is used to cancel self-interference (16QAM, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$). 92

Figure 4.5 Complementary cumulative distribution functions of the maximum achievable sum rate of a multi-user MIMO channel with $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$. SNR = (a) 20 dB and (b) 30 dB. 94

Figure 4.6 Average maximum achievable sum rate of different multi-user MIMO channels ($M = N = 4$, 6, 8, 10, and $m_k = 2$). SNR = (a) 20 dB and (b) 30 dB. 95

Figure 4.7 Structure of the ZF/MMSE nonlinear joint Tx-Rx processing algorithm. 98

Figure 4.8 Structure of the ZF nonlinear joint Tx-Rx processing algorithm when order L is used. 108

Figure 4.9 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $N = 6$, $K = 2$, and $m_1 = m_2 = 3$). 115

Figure 4.10 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $N = 8$, $K = 3$, $m_1 = 2$, and $m_2 = m_3 = 3$). 116

Figure 4.11 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the L-DECOM technique based algorithms (QPSK, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$). 117

Figure 4.12 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the NL-DECOM technique based algorithms (16QAM, $N = 9$, $K = 3$, and $m_1 = m_2 = m_3 = 3$). 118

Figure 4.13 Performance of the MMSE nonlinear joint Tx-Rx processing algorithm (QPSK, $N = 6$, $K = 2$, and $m_1 = m_2 = 3$). 128

Figure 4.14 Performance of the MMSE nonlinear joint Tx-Rx processing algorithm (QPSK, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$). 129

Figure 4.15 Performance comparison of the MMSE nonlinear joint Tx-Rx processing algorithm when no ordering or the "best-first" ordering is used (QPSK, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$). 130

List of Abbreviations

BER	bit error rate
BS	base station
CCDF	complementary cumulative distribution function
CDMA	code division multiple access
CODBF	combined optimal diversity and "best-first"
CSI	channel state information
CSIT	channel state information at the transmitter
DFD	decision feedback detector
DSL	digital subscriber line
GMD	geometric mean decomposition
i.i.d.	independent and identically distributed
ISI	inter-symbol interference
ITU-R	International Telecommunication Union Radiocommunication Sector
L-DECOM	linear multi-user MIMO decomposition
LP	linear pre-processing
LD	linear detector
MIMO	multiple input multiple output
MISO	multiple input single output
MLD	maximum likelihood detector
MMSE	minimum mean-square error
MS	mobile station
MSE	mean-square error
MUD	multi-user detection
MUI	multi-user interference
NL-DECOM	nonlinear multi-user MIMO decomposition

OFDM	orthogonal frequency division multiplexing
SDM	space division multiplexing
SIC	successive interference cancellation
SIMO	single input multiple output
SINR	signal-to-interference plus noise ratio
SISO	single input single output
SNR	signal-to-noise ratio
STBC	space-time block codes
STC	space-time coding
STTC	space-time trellis codes
SVD	singular value decomposition
TDD	time division duplex
THP	Tomlinson-Harashima precoding
TH precoder	Tomlinson-Harashima precoder
Tx-Rx	transmitter-receiver
V-BLAST	vertical Bell Labs layered space-time
ZF	zero-forcing

Chapter 1 Introduction

With the introduction of mobile Internet and mobile multimedia services, the demand for high data rate transmission in wireless communications is rapidly growing. It is stated by International Telecommunication Union Radiocommunication Sector (ITU-R) that up to approximately 100 Mbits/s for mobile access and up to approximately 1 Gbits/s for nomadic and local wireless access are the targets of the future wireless communication systems [1]. However, the radio spectrum that wireless communications can use is a precious and scarce resource. Therefore, techniques that can increase spectral efficiency are of critical importance to the next generation wireless communications.

The multiple input multiple output (MIMO) antenna techniques are among the most prominent breakthroughs in wireless communications in the recent years. By creatively taking advantage of the multipath, they increase the number of degrees of freedom in the channel in the spatial domain, resulting in the capacity increase without bandwidth expansion or transmitted power increase. Since its outset, MIMO antenna techniques have gained considerable research interest and there is a general consensus that they will be a major factor in the future evolution of high data rate wireless systems [1]–[4].

1.1 Brief Introduction to MIMO Systems

In the traditional antenna systems, only the base station is equipped with multiple antennas since the extra cost and space are more affordable at base stations than at user terminals. The uplink, where the transmission is from a mobile to a base station, can be viewed as a single input multiple output (SIMO) antenna system. The downlink, where the transmission is from a base station to a mobile, can be viewed as a multiple input single output (MISO) antenna system. Antenna diversity [5], [6] and beamforming [7], [8] techniques can be applied in the SIMO

and MISO antenna systems. Using antenna diversity techniques, one can transmit the same signal over several independent or highly uncorrelated paths created by different transmit-receive antenna pairs. Thus the probability that the signal over each path undergoes a deep fade at the same time is reduced significantly and hence the reliability of the wireless link is increased. Antenna beamforming techniques, on the other hand, use antenna array processing algorithms to maximize antenna gain in the direction of the desired user. Antenna gains in the directions of the interfering users may also be minimized at the same time. Accordingly, the system performance will be improved as a result of the increased signal-to-noise ratio (SNR) or signal-to-interference plus noise ratio (SINR).

Currently, MIMO antenna systems, where both the transmitter and the receiver are equipped with multiple antennas, are made possible due to hardware advances and relaxed size constraints of some user devices (such as laptops) in high data rate wireless communications. A typical $M \times N$ MIMO system is shown in Figure 1.1. In the transmitter, the data stream to be transmitted is coded, modulated, split into parallel streams, and then mapped onto N transmit antennas. The receiver employs M receive antennas to receive and then recover the transmitted data. If frequency non-selective (flat) fading is assumed and channel is assumed to be constant over each symbol interval, the channel between N transmit antennas and M receive antennas can be represented by an $M \times N$ channel gain matrix $\mathbf{H}$ whose ith row and jth column element, h_{ij}, denotes the complex gain of the channel between the ith receive antenna and the jth transmit antenna.

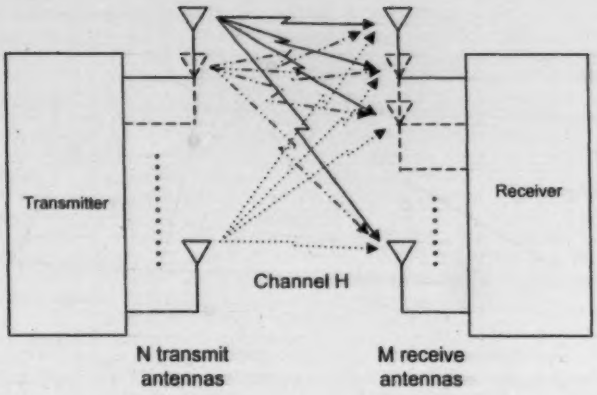


Figure 1.1 Basic structure of an $M \times N$ MIMO system.

Different from the traditional diversity and beamforming techniques, a novel and important idea behind the MIMO techniques is that it makes good use of the multipath, which was previously thought of as a disadvantage in wireless communications, to improve wireless communications system capacity without extra radio spectrum resources or increased transmitted power. It is shown in [9]–[11] that in a rich scattering wireless environment (for example, indoor wireless communications), where the signals arriving at each receive antenna from each transmit antenna undergo highly uncorrelated fading, the MIMO system capacity can be substantially improved compared with the traditional communication systems. Particularly, if the paths between individual transmit-receive antenna pairs are independent and identically distributed (i.i.d.) Rayleigh faded, at high SNR and with the knowledge of channel matrix **H** at the receiver, the capacity of an $M \times N$ MIMO system can be written as

$$C = \min(M, N)\log_2 \mathrm{SNR} + O(1) \quad \text{bits/s/Hz} \tag{1.1}$$

where $\min(M, N)\log_2 \mathrm{SNR}$ is the dominant term and $O(1)$ is a small term of negligible impact in comparison to the SNR [9]–[12]. Eq. (1.1) shows that at high SNR the capacity of an $M \times N$ MIMO system grows linearly with min(M,N). In other words, the capacity of the MIMO channel can be approximately equivalent to the sum capacity of min(M,N) single input single output (SISO) channels.

Researchers have designed different schemes to achieve high spectral efficiency in MIMO systems. Generally, these schemes can be divided into two categories, schemes aiming to exploit the degrees of freedom that a MIMO system can provide, and schemes aiming to exploit the maximum diversity gain that a MIMO system can provide [13]. The tradeoff considerations between these two goals have been investigated in the literature [12], [14], [15].

MIMO systems can provide a gain with respect to the degrees of freedom in the spatial domain. The degrees of freedom, which are defined as the dimensions of the received signal space, dictate the number of different signals that can be reliably distinguished at the receiver [16]. For an $M \times N$ MIMO system, if the paths between individual transmit-receive antenna pairs are i.i.d. Rayleigh faded, the number of spatial degrees of freedom of this system is $\min(M,N)$ [16], which means $\min(M,N)$ independent streams of data can be transmitted through this channel simultaneously. It should be noted that MISO system, SIMO system, and SISO system can only offer a spatial degree-of-freedom of one.

Utilizing the degree-of-freedom gain provided by the MIMO systems, the spatial multiplexing scheme transmits independent streams of data in parallel through the MIMO channel to maximize the transmission data rate. However, the parallel transmission of independent data streams introduces severe inter-stream interference at the receiver end. Therefore, signal processing (spatial layer separation) algorithms at the receiver or/and at the transmitter are required to suppress this inter-stream interference. According to the location of the processing, the signal processing algorithms can be classified as the receiver-processing (or signal detection) algorithms, transmitter pre-processing algorithms, and the joint transmitter-receiver (joint Tx-Rx) processing algorithms.

A typical example of a system that uses spatial multiplexing over MIMO channel, the vertical Bell Labs layered space-time (V-BLAST) [17], [18], demonstrates very high spectral efficiency. Unprecedented wireless spectral efficiencies of 20-40 bits/s/Hz in a MIMO system with $N = 8$ transmit and $M = 12$ receive antennas have been observed. This result shows that the MIMO systems using spatial multiplexing have a big potential in commercial applications.

Besides the degree-of-freedom gain, there is also a diversity gain in the MIMO systems. The diversity gain determines the slope at which the average error probability decays at high SNR. That is, if d is the diversity gain, the average error probability decays as $1/\mathrm{SNR}^d$. The maximum diversity gain of an $M \times N$ MIMO system is MN. It should be noted that the diversity gain also exists in an $M \times 1$ SIMO system and a $1 \times N$ MISO system where the maximum diversity gains are M and N, respectively. Space-time coding (STC) schemes are used to achieve the maximum diversity gain for MIMO systems. In a STC scheme, a space-time encoder at the transmitter encodes the data and outputs N streams, which will then be transmitted by N transmit antennas. The space-time codes can be classified as the space-time trellis codes (STTC) [19] and the space-time block codes (STBC) [20], [21]. The STTC can provide the full diversity gain MN, but it requires trellis decoding at the receiver and may imply very high decoding complexity [19]. The most significant benefit of the STBC is that the decoding uses linear processing, which makes it much simpler than the decoding of the STTC. The STBC can also achieve full diversity gain. It should be noted that unlike in spatial multiplexing, on average at most one independent stream of data could be transmitted through the MIMO channel using the STTC and the STBC.

1.2 Brief Introduction to Multi-User MIMO Systems

Most recently, multi-user MIMO systems have been given much attention since the accommodation of multiple users is necessary in public wireless communication systems. The multi-user MIMO system is defined as a system in which multiple mobile stations with one or more antennas are simultaneously communicating with the base station equipped with multiple antennas. Data transmission of different users over the same channel is achieved by space division multiplexing (SDM).

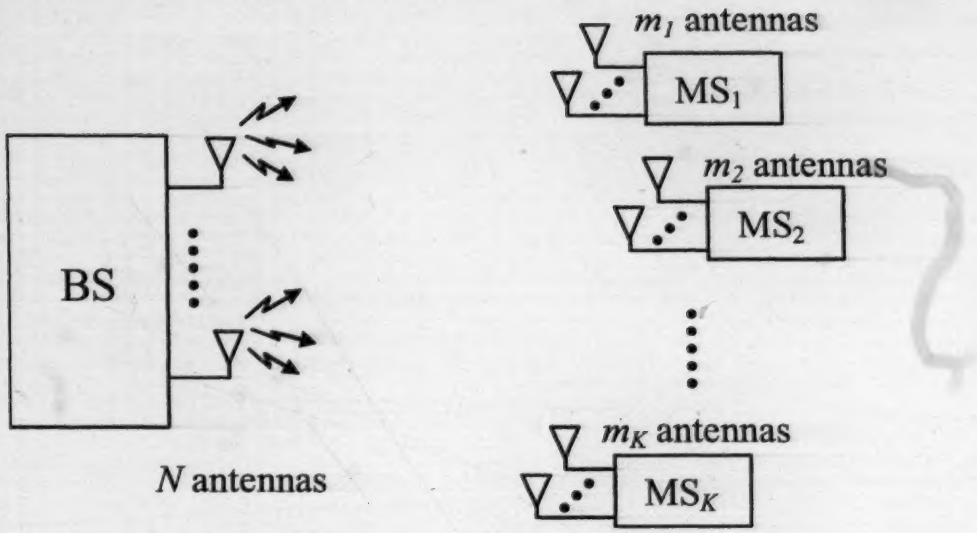


Figure 1.2 Basic structure of a multi-user MIMO system.

Figure 1.2 shows the basic structure of a multi-user MIMO system. Here we assume that the base station (BS) has N antennas and there are K mobile stations (MSs), MS_1, MS_2, ... and MS_K, in this system. The mobile station MS_i has m_i antennas, and $M = \sum_{i=1}^{K} m_i$ is the total number of antennas at all mobiles. This system can be seen as a point-to-point MIMO antenna system with N antennas at one end and M antennas at the other end, but the mobiles cannot process signals cooperatively. If the channel paths between individual transmit-receive antenna pairs are i.i.d. Rayleigh faded, it is shown that at high SNR the sum capacity (total transmission rate of all mobiles) also grows linearly with $\min(M,N)$ for the uplink provided that the base station knows the channel state information (CSI) [10], [22], and for the downlink provided that the base station and the mobiles all know the CSI [16], [23]–[26].

1.3 Precoding for Spatial Layer Separation on Multi-user MIMO Downlink and Summary of Contributions

In multi-user MIMO systems, through spatial multiplexing, multiple data streams can be transmitted within the same frequency band and in the same time slot. Thus very high throughput can be achieved. However, the parallel transmission of independent data streams introduces severe inter-stream interference at the receiver end. Therefore, the cancellation of the inter-stream interference (or the spatial layer separation) is an important problem for the multi-user MIMO systems.

In the uplink of multi-user MIMO systems, a large body of receiver-processing (signal detection) algorithms developed for single-user MIMO systems can be applied at the base station to recover the transmitted data vector from the received signal vector. The V-BLAST algorithm [17], [18], [27], [28] is a well-known example of this approach. An iterative "nulling and cancellation" method is used in the V-BLAST, and this method can be shown to be equivalent to the decision-feedback detector (DFD) in [29]. The maximum likelihood detector (MLD) [30]–[32], the zero-forcing (ZF) and minimum mean-square error (MMSE) linear detector [30] are other examples of signal detection algorithms that can be used to cancel the inter-stream interference. These algorithms are all fundamentally the same as multi-user detection (MUD) algorithms that have been proposed for code division multiple access (CDMA) systems. There are also many other signal detection algorithms with various performance and complexity tradeoffs discussed in the literature, e.g., [33]–[39].

However, in the multi-user MIMO downlink, receiver-processing algorithms cannot be used due to the absence of coordination among receivers of different independent mobile users. Pre-processing algorithms have to be applied at the transmitter to pre-distort the signals of multiple users in such a way that they no longer interfere at the individual antennas of mobile receivers. Compared to the well-developed receiver-processing algorithms, there are many open questions for the transmitter pre-processing algorithms.

Linear transmitter pre-processing algorithms based on the ZF and MMSE criteria have been proposed in [40]–[43], but their performance is only comparable to that of the linear detection algorithms using the same criterion (ZF or MMSE) since they suffer from the transmitted power increase [43], [44]. Therefore, the design of the nonlinear pre-processing algorithms, which can achieve better performance, has attracted increasing research interests recently.

The optimal capacity-maximizing nonlinear precoding is the Costa's dirty-paper coding [45]. Costa proves that for a SISO channel if interference is known at the transmitter but not known at the receiver, the full non-causal knowledge of the interference can be used in the precoding so that the same capacity as if the interference is absent can be achieved. This result is utilized in [24] to achieve the sum capacity of the multi-user MIMO downlink. The scheme in [24] applies linear precoding to make the channel spatially causal, and then the dirty-paper coding is used to cancel the residual causal inter-stream interference. However, the optimal dirty-paper coding is not a practical scheme, since it operates on blocks whose lengths approach infinity. A lattice quantization based precoding strategy proposed in [46] and [47] was shown to be able to achieve the same system capacity as the dirty-paper coding, and this strategy can also be used to derive the sum capacity of the multi-user MIMO downlink. However, the dimension of the optimum lattice quantizers used in this strategy approaches infinity, which also prohibits its practical implementation. The Tomlinson-Harashima precoding (THP) technique [48]–[50] first proposed to pre-equalize the inter-symbol interference (ISI) in a dispersive SISO channel, can be seen as a practical implementation of the dirty-paper coding. It applies a modulo device to constrain the transmitted power increase, so better performance can be achieved as compared to the linear precoding techniques [51]. The application of the THP for spatial layer separation in the multi-user MIMO downlink has been recently proposed by [51] and [52]. Since the design of implementable precoding algorithms is the focus of this thesis, nonlinear pre-processing algorithms utilizing the THP are of our interest.

It has been shown that the nonlinear pre-processing algorithm of [51] and [52] can achieve significantly better performance than the linear pre-processing algorithm when used in the multi-user MIMO downlink [52]. The processing matrices of the nonlinear pre-processing algorithm of [51] and [52] satisfy the ZF criterion, and the algorithm of [51] and [52] is called the conventional ZF nonlinear pre-processing algorithm in this thesis. However, for the nonlinear pre-processing structure proposed in [51] and [52], the solution for the processing matrices that satisfies the ZF criterion is not unique. The processing matrices of the conventional ZF nonlinear pre-processing algorithm are introduced intuitively without any optimization. Also, the design of nonlinear pre-processing algorithm based on the MMSE criterion needs to be analyzed. Therefore, this thesis addresses the above questions and makes the following contributions.

- We analyze possible solutions for the processing matrices that satisfy the ZF criterion for the nonlinear pre-processing in a multi-user MIMO downlink, and then propose a generalized ZF nonlinear pre-processing structure. This structure can realize ZF-based nonlinear pre-processing algorithm that achieves the minimum bit error rate (BER) for each mobile under any given power allocation to different data streams. The choice of power allocation scheme for the generalized ZF nonlinear pre-processing structure based on different performance requirements is also discussed. Based on this structure, we have designed the ZF nonlinear pre-processing algorithm, which achieves minimum BER at each mobile under any given relative SNR requirement, the ZF nonlinear pre-processing algorithm, which achieves minimum total transmitted power while satisfying the individual SNR target at each mobile, and the minimum BER ZF nonlinear pre-processing algorithm, which achieves minimum average BER of the system.
- An MMSE criterion based nonlinear pre-processing algorithm, the MMSE nonlinear pre-processing algorithm, is proposed for multi-user MIMO downlinks. This algorithm can achieve significantly better performance than its ZF criterion based counterpart.

For the conventional ZF nonlinear pre-processing algorithm of [51] and [52] and the MMSE nonlinear pre-processing algorithm we propose, it can be found that the system performance can be improved if the rows of the channel matrix are suitably ordered. We present solutions of how to find a proper order for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm in this thesis. Our contributions can be summarized as follows:

- An optimal ordering lemma is proposed. This lemma gives the conditions under which the "best-first" ordering method in the V-BLAST systems [18], [28] can achieve the optimal order.
- The optimal ordering lemma is subsequently used to solve the ordering problems in the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm. Using this lemma, it is shown that the "best-first" ordering method can achieve the optimal order in the minimax noise variance sense for the conventional ZF nonlinear pre-processing algorithm and the optimal order in the minimax error variance sense for the MMSE nonlinear pre-processing algorithm.
- Other applications of the optimal ordering lemma will also be addressed in this thesis.

For the downlink of multi-user MIMO systems with multiple-antenna mobiles, it is possible to utilize the processing capabilities of the mobiles to design joint Tx-Rx processing algorithms to achieve better performance than with pre-processing only algorithms. Linear joint Tx-Rx processing algorithms, where linear transmitter processing and linear receiver processing are applied at the base station and each mobile, have been proposed in [53]–[56] for the downlink of multi-user MIMO systems with multiple-antenna mobiles, but the processing matrices there have to be found by iterative methods, which leads to high computational complexity. In [57]–[60], a linear multi-user MIMO decomposition (L-DECOM) technique, which realizes joint Tx-Rx processing and provides closed-form solution, is proposed. This technique uses linear pre-processing to

pre-eliminate multi-user interference (MUI), the interference from the data streams of other users, at the transmitter. The data streams of each user can be seen as passing through an equivalent single-user MIMO channel without any MUI. The self-interference from other data streams of the same user can then be mitigated by any single-user MIMO layer separation algorithm. With the application of different single-user MIMO processing algorithms, L-DECOM technique based joint Tx-Rx processing algorithms for the multi-user MIMO downlink with multiple-antenna mobiles are realized, and closed-form expressions for processing matrices can be derived. However, the performance of these joint Tx-Rx processing algorithms is limited by the linear pre-processing used by the L-DECOM technique. The design of the nonlinear joint Tx-Rx processing algorithms where nonlinear THP is applied to constrain the transmitted power increase is still an open question. This thesis will also investigate this question and our contributions in this regard include:

- A nonlinear multi-user MIMO decomposition (NL-DECOM) technique for the multi-user MIMO downlink with multiple-antenna mobiles. This technique uses nonlinear pre-processing instead of linear pre-processing to pre-eliminate the MUI at the transmitter. When this technique is applied the data streams of each user can also be seen as passing through an equivalent single-user MIMO channel without any MUI, as is the case with the L-DECOM technique. It will be shown that the NL-DECOM technique achieves better BER performance, realizes much higher maximum achievable sum rate at high SNR, and affords higher flexibility in system design than the L-DECOM technique.
- The nonlinear pre-processing in the NL-DECOM technique only pre-eliminates the MUI. When the NL-DECOM technique is used, since the nonlinear pre-processing at the transmitter changes the transmitted signal [51], the single-user MIMO spatial layer separation algorithms that employ the THP cannot be applied directly to cancel the self-interference. On the other hand, for the joint Tx-Rx processing of

multi-user MIMO downlink with multiple-antenna mobiles, to achieve better performance and reduce the complexity of the mobiles, it is desirable to introduce algorithms that apply nonlinear THP to suppress both the MUI and self-interference at the transmitter. In this thesis we propose the ZF nonlinear joint Tx-Rx processing algorithm and the MMSE nonlinear joint Tx-Rx processing algorithm. In these algorithms, nonlinear THP and linear receiver-processing are used to perform the spatial layer separation on multi-user MIMO downlink with multiple-antenna mobiles. These joint Tx-Rx algorithms effectively utilize the processing capabilities of the base station and the mobiles, while the complexity of the mobiles is kept low since linear receiver processing is applied.

1.4 Thesis Organization

Chapter 2 focuses on the design of nonlinear pre-processing algorithms for spatial layer separation on the multi-user MIMO downlink. First, system model and assumptions will be described. As background knowledge, the principle of operation of the THP will be introduced, where we will show how the modulo device in the THP can constrain the transmitted power. The conventional ZF nonlinear pre-processing algorithm of [51] and [52] in which the processing matrices are introduced intuitively without any optimization will be briefly introduced. Then, a generalized ZF nonlinear pre-processing structure will be proposed in this chapter. This structure can realize ZF-based nonlinear pre-processing algorithm that achieves minimum BER for each mobile under any given power allocation to different data streams. Then, the choice of power allocation scheme based on different performance optimization requirements will be discussed. We will show that the conventional ZF nonlinear pre-processing algorithm can be seen as a special implementation of the generalized ZF nonlinear pre-processing structure. In this chapter, the nonlinear pre-processing algorithm based on the MMSE criterion (the MMSE nonlinear pre-processing algorithm) will also be proposed for the downlink of multi-user MIMO system. This

algorithm achieves significantly better performance than the ZF criterion based counterpart since it avoids the noise enhancement effect by minimizing the mean-square error (MSE).

In Chapter 3, the ordering problems of the nonlinear pre-processing algorithms employing the THP will be studied. We will first show the importance of ordering in the nonlinear pre-processing algorithms employing the THP, and it will be found that if the rows of the channel matrix are suitably ordered, better performance can be achieved. The change of the system structure caused by ordering the rows of the channel matrix will be revealed. Then, an optimal ordering lemma will be introduced. This lemma gives the conditions under which the "best-first" ordering method is optimal. Using this lemma, we will show that the "best-first" ordering method can achieve the optimal order in the minimax noise variance sense for the conventional ZF nonlinear pre-processing algorithm and the optimal order in the minimax error variance sense for the MMSE nonlinear pre-processing algorithm we propose. An efficient way to perform the "best-first" ordering for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm will be presented. The ordering problems of the generalized ZF nonlinear pre-processing structure based nonlinear pre-processing algorithms proposed in Chapter 2 will also be addressed.

Chapter 4 is concerned with the design of nonlinear joint Tx-Rx processing algorithms for the downlink transmission of multi-user MIMO systems where the mobiles are equipped with multiple antennas. The system model and assumptions for the multi-user MIMO downlink with multiple-antenna mobiles will first be described. The L-DECOM technique [57]–[60], which realizes the decomposition of the multi-user MIMO channel into parallel independent single-user MIMO channels by linear pre-processing, will then be introduced. Our NL-DECOM technique for the multi-user MIMO downlink with multiple-antenna mobiles will be proposed, and the advantages of the NL-DECOM technique over the L-DECOM technique will be analyzed. It will be shown that the NL-DECOM technique can achieve better BER performance and achieve much higher maximum achievable sum rate at high SNR than the L-DECOM technique.

Moreover, the restriction on the number of transmit antennas and the restriction on the number of data streams for each mobile in the NL-DECOM technique are less strict than those in the L-DECOM technique, which means the NL-DECOM technique affords higher flexibility in system design. When the NL-DECOM technique is applied, it will be found that the mobiles can be ordered properly to further improve the system performance, and the "best-first" ordering method is proposed to improve the overall system performance in the systems, in which all the mobiles have the same number of antennas and the same self-interference cancellation algorithm is used for each mobile. In order to overcome the problem that the single-user MIMO layer separation algorithms employing the THP cannot be combined with the NL-DECOM technique, the ZF and MMSE nonlinear joint Tx-Rx processing algorithms will be proposed. A ZF criterion based algorithm, the ZF nonlinear joint Tx-Rx processing algorithm, can ensure the complete elimination of the inter-stream interference, while it does not require the knowledge of the statistics of the transmitted signal and the noise. An MMSE criterion based algorithm, the MMSE nonlinear joint Tx-Rx processing algorithm, minimizes the MSE, which includes both the residual inter-stream interference and noise. Hence, it mitigates noise enhancement, and achieves better performance in comparison with its ZF-based counterpart. In these two algorithms, closed-form expressions for the transmitter and receiver processing matrices are developed to optimize the performance of each mobile.

Chapter 5 concludes this thesis by summarizing the major contributions and suggesting future directions of research work.

1.5 Notation

Some symbols and operators used in this thesis are defined here. Boldface small symbol denotes a vector (e.g., $\mathbf{a}$), and boldface capital symbol denotes a matrix (e.g., $\mathbf{H}$). Notation diag$[x_1,x_2,\ldots,x_D]$ denotes a diagonal matrix with diagonal elements $x_1, x_2, \ldots, x_D$, while diag($\mathbf{A}$) denotes a vector made up of the diagonal elements of $\mathbf{A}$. The null space [61] of $\mathbf{A}$ is denoted by null($\mathbf{A}$), and the range [61] of $\mathbf{A}$ is denoted by ran($\mathbf{A}$). $[\mathbf{A}]_{ij}$ denotes the element at the ith row and jth column

of $\mathbf{A}$, $(\mathbf{A})_i$ denotes the ith row of $\mathbf{A}$, and $[\mathbf{A}]_j$ denotes the jth column of $\mathbf{A}$. $\mathbf{0}_{x,y}$ denotes an $x \times y$ matrix with all zero elements. The real part of a is denoted by $\mathrm{Re}\{a\}$, and the imaginary part of a is denoted by $\mathrm{Im}\{a\}$. $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. $\mathrm{E}[\cdot]$ denotes expectation, $(\cdot)^T$ denotes a transpose, while $(\cdot)^H$ denotes a conjugate transpose. $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse (Assuming $\mathbf{A}$ is an $M \times N$ matrix, $\mathbf{A}^\dagger$ is: if $M \leq N$, $\mathbf{A}^\dagger = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$; if $M = N$, $\mathbf{A}^\dagger = \mathbf{A}^{-1}$; if $M \geq N$, $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$) [61].

Chapter 2 Nonlinear Pre-Processing Algorithms for Multi-User MIMO Downlink

In this chapter, the design of nonlinear pre-processing algorithms for spatial layer separation in the multi-user MIMO downlink will be considered.

To realize spatial layer separation in the multi-user MIMO downlink, linear transmitter pre-processing algorithms proposed in [40]–[43] can be applied. However, these algorithms cause significant transmitted power increase, which greatly limits their performance [62]. As a result, the performance of the linear pre-processing algorithms is only comparable to that of the linear detection algorithms, and hence there is a large performance gap compared with nonlinear detection algorithms [43], [44]. In [51] and [52], a nonlinear pre-processing algorithm was proposed. This algorithm utilizes the THP device to constrain the transmitted power increase while pre-eliminating the inter-stream interference. Thus, it achieves significantly better performance than the linear pre-processing algorithms, and its performance is comparable to or even better than that of the nonlinear DFD algorithm due to the avoidance of the error propagation [44], [52]. The processing matrices of the nonlinear pre-processing algorithm of [51] and [52] satisfy the ZF criterion, so we call this algorithm the conventional ZF nonlinear pre-processing algorithm in this thesis.

For the nonlinear pre-processing structure proposed in [51] and [52], the solution of the processing matrices that satisfies the ZF criterion is not unique. The processing matrices of the conventional ZF nonlinear pre-processing algorithm are introduced intuitively without any optimization. In this chapter, a

generalized ZF nonlinear pre-processing structure is proposed. This structure can realize ZF-based nonlinear pre-processing algorithm that achieves minimum BER for each mobile under any given power allocation to different data streams. Then, the choice of power allocation scheme based on different performance requirements is discussed. Based on this structure, we have designed the ZF nonlinear pre-processing algorithm, which achieves minimum BER at each mobile under any given relative SNR requirement, the ZF nonlinear pre-processing algorithm, which achieves minimum total transmitted power, while satisfying the individual SNR target at each mobile, and the minimum BER ZF nonlinear pre-processing algorithm, which achieves minimum average BER of the system. Moreover, it will be shown that the conventional ZF nonlinear pre-processing algorithm can be seen as a special implementation of the generalized ZF nonlinear pre-processing structure, and the generalized ZF nonlinear pre-processing structure provides a simplified way to achieve power allocation for the conventional ZF nonlinear pre-processing algorithm.

On the other hand, MMSE criterion can lead to better performance than the ZF criterion since the noise enhancement effect can be avoided [44], [63], [64]. Therefore, design of the nonlinear pre-processing algorithm based on the MMSE criterion becomes of interest. Nonlinear pre-processing algorithms that minimize the total transmitted power while satisfying individual MSE (or SINR) target at each mobile have been proposed in [16], [65], and [66], based on the uplink-downlink duality between linear precoding (downlink beamforming) combined with dirty-paper coding and linear receiver processing (uplink beamforming) combined with perfect successive interference cancellation (SIC) discussed in [16], [26], [65], and [66]. The optimum solutions in these algorithms have to be found by iterative methods, which makes their computational complexity high. In this thesis, the MMSE-criterion-based nonlinear pre-processing algorithm which minimizes the sum MSE at all the receive branches (the trace of the error correlation matrix) under a fixed total transmitted power is proposed. This algorithm achieves significantly better performance than the ZF criterion based counterpart. Moreover, a closed-form expression for the processing matrices is

found for the proposed algorithm, which makes it potentially attractive for practical implementation.

In the following, system model and assumptions will first be described in Section 2.1. As background knowledge, the principle of operation of the THP will be briefly introduced in Section 2.2. The conventional ZF nonlinear pre-processing algorithm of [51] and [52] will be described in Section 2.3. In Section 2.4, the generalized ZF nonlinear pre-processing structure will be proposed. Finally, the MMSE nonlinear pre-processing algorithm will be proposed in Section 2.5.

2.1 System Model and Assumptions

In this thesis, the downlink transmission of multi-user MIMO systems is considered. Flat Rayleigh fading channel is assumed. The channel is also assumed to be constant over each symbol interval and change independently from one symbol interval to the next. We assume the transmitter has perfect knowledge of the CSI. The knowledge of the channel state information at the transmitter (CSIT) can be obtained by channel estimation in time division duplex (TDD) systems (such as IEEE 802.16e [67], HIPERLAN/2 [68], and TD-SCDMA/UTRA TDD 1.28 Mcps option [69]) where the uplink and downlink are reciprocal or by feedback from the receivers.

Since pre-processing algorithms are the focus of this chapter, for simplicity we assume that only one antenna is equipped at each mobile. It is noted that the algorithms discussed and proposed in this chapter can be utilized directly in the multi-user MIMO systems where some or all the mobiles are equipped with multiple antennas.

It is assumed that there are N transmit antennas, which are denoted as $\text{Ant}_1,\text{Ant}_2,\ldots,\text{Ant}_N$, deployed at the base station. There are M mobiles simultaneously communicating with the base station, and the mobiles are denoted as $\text{MS}_1,\text{MS}_2,\ldots,\text{MS}_M$. Vector $\mathbf{a} \equiv (a_1, a_2, ..., a_M)^T$ is the data vector, where data symbol a_i is for MS_i. It is assumed that all the data symbols are independent and have unit power, i.e., $\mathbf{R}_{\mathbf{a}} \equiv \mathrm{E}[\mathbf{a}\mathbf{a}^H] = \mathbf{I}$. To ensure the transmission of M

independent streams of data in this system, the number of transmit antennas should be larger than or equal to M, so $N \geq M$ is assumed[1].

Since pre-processing is applied, we assume $\mathbf{x} \equiv (x_1, x_2, ..., x_N)^T$ is the pre-processing output, where $x_1, x_2, ..., x_N$ are the signals transmitted by the transmit antennas $\text{Ant}_1, \text{Ant}_2, \ldots, \text{Ant}_N$. The received signal at MS_i is

$$y_i = (\mathbf{H})_i \mathbf{x} + n_i \tag{2.1}$$

where $(\mathbf{H})_i \equiv (h_{i1}, h_{i2}, ..., h_{iN})$ is a row vector made up of the channel gains from the Ant_1, $\text{Ant}_2, \ldots, \text{Ant}_N$ to MS_i, and n_i is the additive white zero-mean complex Gaussian noise at the input of MS_i.

The received signals at all the mobiles can be written into one larger vector

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.2}$$

where $\mathbf{y} \equiv (y_1, y_2, ..., y_M)^T$, $\mathbf{n} \equiv (n_1, n_2, ..., n_M)^T$, and $\mathbf{H} \equiv [(\mathbf{H})_1^T, (\mathbf{H})_2^T, ..., (\mathbf{H})_M^T]^T$.

Since flat Rayleigh fading is assumed, the entries of $\mathbf{H}$ are assumed to be normalized i.i.d. zero-mean complex Gaussian random variables. $n_1, n_2, ..., n_M$ are assumed to be mutually independent and have the same variance, i.e., $\mathbf{R}_n \equiv \text{E}[\mathbf{n}\mathbf{n}^H] = \sigma_n^2 \mathbf{I}$.

In this thesis, the decision device applies symbol-by-symbol maximum likelihood detection. Channel coding is not considered in this thesis, since our focus is on the spatial layer separation algorithms. It should be noted that by encoding/decoding the independent data streams individually, channel coding can be effectively integrated into the spatial multiplexing scheme. We also assume that all the data symbols are using the same modulation size.

---

[1] In systems with a large number of mobiles, this assumption can be satisfied by mobile scheduling schemes, which select a subset of mobiles for communication with the base station simultaneously. The mobile scheduling schemes are beyond the scope of this thesis, and they are one direction of possible future research work.

2.2 Introduction to the Tomlinson-Harashima Precoder

The Tomlinson-Harashima (TH) precoder was originally proposed to pre-equalize the ISI in a dispersive SISO channel in [48]–[50]. Its structure is shown in Figure 2.1.

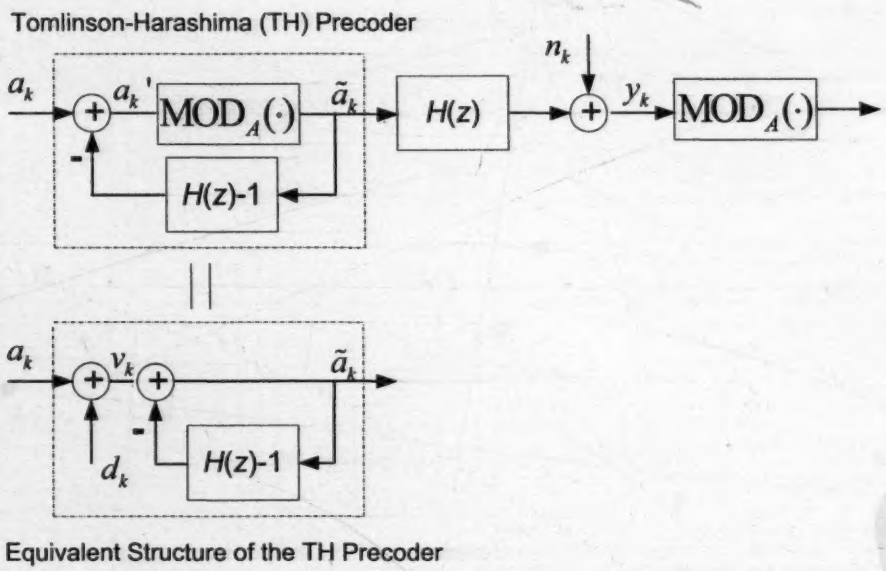


Figure 2.1 Structure of the TH precoder for SISO channel with ISI.

In Figure 2.1, a_k is the signal to be transmitted, and n_k is the noise sample at the receiver. $H(z)$ is the z -transform of the channel impulse response, i.e., $H(z)=\sum_k h_k z^{-k}$. The channel is assumed to have L taps while $h_0=1$. If precoding is not applied, the received signal can be represented as $y_k = a_k + \sum_{i=1}^{L-1} h_i a_{k-i} + n_k$. The term $\sum_{i=1}^{L-1} h_i a_{k-i}$ is the ISI for the signal a_k . If the channel is known at transmitter, the ISI can be simply pre-eliminated by transmitting a precoded signal $a_k' = a_k - \sum_{i=1}^{L-1} h_i a_{k-i}'$ instead of a_k . However, this will lead to an increase of the transmitted power. To avoid the transmitted power

increase, in the TH precoder the precoded signal a_k' goes through a modulo device $\mathrm{MOD}_A(\cdot)$ before it is launched into the channel. The function of the modulo device $\mathrm{MOD}_A(\cdot)$ is to constrain its output into $[-A/2, A/2)$, i.e.,

$$\tilde{y} = \mathrm{MOD}_A(y) = y - A\lfloor (y + A/2)/A \rfloor = y + d \tag{2.3}$$

where operation $\lfloor x \rfloor$ finds the nearest integer less than or equal to x and $d = -A\lfloor (y + A/2)/A \rfloor$ is an unique integer multiple of A. The value of A is determined according to the power of the transmitted signal and the modulation used [51]. For example, if M_c -ary PAM is used and a_i is from the set $\{\pm 1, \pm 3, ..., \pm(M_c - 1)\}$, A should be chosen as $2M_c$. If M_c -ary square QAM is used and a_i is from the set $\{a^I + ja^Q \mid a^I, a^Q \in \{\pm 1, \pm 3, ..., \pm(\sqrt{M_c} - 1)\}\}$, A should be chosen as $2\sqrt{M_c}$.

For the system shown in Figure 2.1, if we assume both the transmitted signal and the channel impulse response are real, the output of the modulo device is

$$\tilde{a}_k = \mathrm{MOD}_A(a_k') = a_k' - A\lfloor (a_k' + A/2)/A \rfloor = a_k' + d_k = a_k + d_k - \sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}. \tag{2.4}$$

Thus, the TH precoder can be seen as generating an effective data symbol $v_k = a_k + d_k$ where $d_k = -A\lfloor (a_k' + A/2)/A \rfloor$ is called precoding symbol. Then, v_k is passed through a feedback filter to pre-eliminate the ISI. It is generated by the modulo device to constrain the dynamic range of the precoding output. This equivalent structure of the TH precoder is also shown in Figure 2.1. This equivalent structure provides a linear representation of the TH precoder, so it is extensively used in the design of processing algorithms employing THP.

Applying THP, one can find that the received signal becomes $y_k = a_k + d_k + n_k$. Since d_k can be removed by the modulo device $\mathrm{MOD}_A(\cdot)$ at the receiver end, the signal a_k is transmitted through the system without any ISI.

In the above introduction of the TH precoder, the transmitted signal and the channel are assumed to be real. If the transmitted signal and the channel are both complex, the modulo device should operate on both the real part and the

imaginary part of its input [51], [70]. We use a_k^I and a_k^Q to represent the real and imaginary parts of a_k, and use $\tilde{a}_k^I$ and $\tilde{a}_k^Q$ to represent the real and imaginary parts of the precoding output $\tilde{a}_k$. Thus, one has

$$\tilde{a}_k^I = \mathrm{MOD}_A(a_k^I - \mathrm{Re}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\}) = a_k^I + d_k^I - \mathrm{Re}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\} \tag{2.5}$$

$$d_k^I = -A\left\lfloor (a_k^I - \mathrm{Re}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\} + A/2)/A \right\rfloor \tag{2.6}$$

$$\tilde{a}_k^Q = \mathrm{MOD}_A(a_k^Q - \mathrm{Im}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\}) = a_k^Q + d_k^Q - \mathrm{Im}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\} \tag{2.7}$$

$$d_k^Q = -A\left\lfloor (a_k^Q - \mathrm{Im}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\} + A/2)/A \right\rfloor \tag{2.8}$$

It can be easily verified that the received signal after going through the modulo device is the ISI free version of the transmitted signal and the equivalent structure of the TH precoder shown in Figure 2.1 can also be used.

Another situation is the application of the THP for real signals on complex channels, which is not explicitly explained in the literature. Transmitting BPSK signals on wireless channels is an example of this scenario. In this situation, since a_k is a real signal, to effectively constrain the transmitted power, the output of the TH precoder should also be a real signal, so the imaginary part of the input to the TH precoder should be discarded. In this case, the linear description of the TH precoder in Figure 2.1 can still be used to represent the precoding procedure, but the real and imaginary parts of the precoding symbol d_k should be generated differently. The real part of d_k is still generated using the modulo device as in (2.6). However, the imaginary part of d_k, d_k^Q, should be a real number (no longer restricted to an integer multiple of A) which lets $\tilde{a}_k^Q = 0$, i.e.,

$$\tilde{a}_k^Q = (a_k^Q - \mathrm{Im}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\}) + d_k^Q = a_k^Q + d_k^Q - \mathrm{Im}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\} = 0 \tag{2.9}$$

$$d_k^Q = -(a_k^Q - \mathrm{Im}\{\sum_{i=1}^{L-1} h_i \tilde{a}_{k-i}\}) \tag{2.10}$$

Therefore, $\tilde{a}_k$ will be a real signal as a_k, and its value is restricted to $[-A/2, A/2)$. At the receiver, the received signal is $y_k = a_k + d_k + n_k$. Only the real part of y_k goes into the modulo and decision device, while the imaginary part of y_k is discarded, so d_k^Q leaves no influence on the decision of a_k.

It should be noted that using the TH precoder, although the value of the precoder output, $\tilde{a}_k$, is restricted to a specific range, since the value of $\tilde{a}_k$ is not taken from the constellation points, there is still a transmitted power increase. However, this transmitted power increase is significantly smaller as compared to that with linear pre-processing algorithms [62], and it is negligible for moderate modulation sizes and vanishes completely as the modulation size approaches infinity [51]. The *precoding loss*, which is defined as $\gamma_p^2 \equiv \mathrm{E}[|\tilde{a}|^2]/\mathrm{E}[|a|^2]$, is used to measure this transmitted power increase [51].

2.3 Conventional ZF Nonlinear Pre-Processing Algorithm

The structure of the conventional ZF nonlinear pre-processing algorithm for the downlink of multi-user MIMO systems proposed in [51] and [52] is shown in Figure 2.2. In this structure, matrix **F** (size $N \times M$) at the transmitter works as a feedforward filter, and matrix (**B-I**) works as a feedback filter, where **B** (size $M \times M$) should be a unit lower triangular matrix (a lower triangular matrix with ones on the main diagonal) [61]. In Figure 2.2, the block outlined by the dashed line at the transmitter is the TH precoder. From the introduction of the TH precoder in Section 2.2, one can see that the precoder equivalently generates an effective data vector $\mathbf{v} \equiv (\mathbf{v}_1^T, \mathbf{v}_2^T, ..., \mathbf{v}_K^T)^T = \mathbf{a} + \mathbf{d}$ as the input to $\mathbf{B}^{-1}$, where $\mathbf{d} \equiv (\mathbf{d}_1^T, \mathbf{d}_2^T, ..., \mathbf{d}_K^T)^T$ is the precoding vector used to constrain the transmitted power. $\tilde{\mathbf{a}} \equiv (\tilde{\mathbf{a}}_1^T, \tilde{\mathbf{a}}_2^T, ..., \tilde{\mathbf{a}}_K^T)^T = \mathbf{B}^{-1}\mathbf{v} = \mathbf{B}^{-1}(\mathbf{a} + \mathbf{d})$ is the precoded output vector.

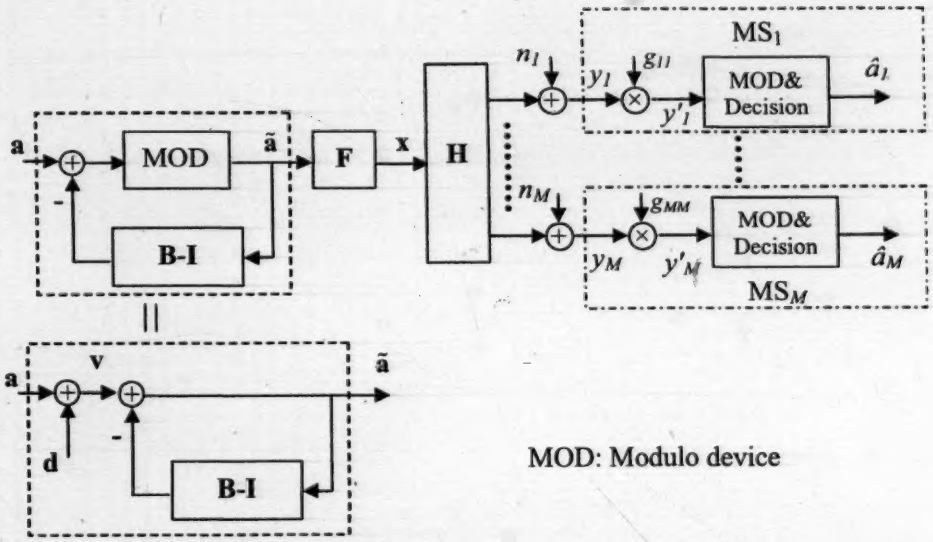


Figure 2.2 Structure of the nonlinear pre-processing on the downlink of a multi-user MIMO system.

At the receiver end, the modulo operation can retrieve the data vector **a** from the received effective data vector **v**. In the structure shown by Figure 2.2, each received signal needs to be scaled before going into the modulo and decision device. The scaling factors at all the mobiles, $g_{11}, g_{22}, \ldots, g_{MM}$, are determined at the transmitter. Hence, their values need to be transmitted to the mobiles.

For the system structure illustrated in Figure 2.2, we introduce a diagonal matrix $\mathbf{G} \equiv \text{diag}[g_{11}, g_{22}, \ldots, g_{MM}]$, whose diagonal elements are made up of the scaling factors at all the receivers. The signals $y'_1, y'_2, \ldots, y'_M$ that will go through the modulo and decision devices can be represented as

$$\begin{aligned} \mathbf{y}' &= \mathbf{G}\mathbf{y} \\ &= \mathbf{GHFB}^{-1}\mathbf{v} + \mathbf{Gn} \\ &= \mathbf{GHFB}^{-1}\mathbf{v} + \tilde{\mathbf{n}} \end{aligned} \tag{2.11}$$

where $\mathbf{y}' \equiv (y'_1, y'_2, \ldots, y'_M)^T$, and $\tilde{\mathbf{n}} \equiv (\tilde{n}_1, \tilde{n}_2, \ldots, \tilde{n}_M)^T = \mathbf{Gn}$.

The ZF criterion ensures that all the inter-stream interference is pre-eliminated, so

$$\mathbf{GHFB}^{-1} = \mathbf{I} \tag{2.12}$$

is required.

For the nonlinear pre-processing structure shown in Figure 2.2, although the power of the TH precoder output $\tilde{\mathbf{a}}$ is constrained by the modulo operation, the transmitted power is still influenced by the linear pre-processing matrix **F**. For a practical system the available transmitted power is fixed, so transmitted power constraint should be applied in the design of the pre-processing matrix **F**. The most commonly used transmitted power constraint is the average transmitted power constraint, which requires the average total transmitted energy per symbol interval $\bar{E}_{tr}$ to be a constant. Since

$$\bar{E}_{tr} = \mathrm{E}[\mathbf{x}^H\mathbf{x}] = \mathrm{tr}\{\mathbf{F}\mathbf{R}_{\tilde{\mathbf{a}}}\mathbf{F}^H\} = \mathrm{tr}\{\mathbf{F}\mathbf{F}^H\}\sigma_{\tilde{\mathbf{a}}}^2, \tag{2.13}$$

constraint on **F** as

$$\mathrm{tr}\{\mathbf{F}\mathbf{F}^H\} = M \tag{2.14}$$

can make the average transmitted power constant. Here the assumption that $\mathbf{R}_{\tilde{\mathbf{a}}} = \sigma_{\tilde{\mathbf{a}}}^2\mathbf{I}$ is used as in [51] and [52], and the value of $\sigma_{\tilde{\mathbf{a}}}^2$ for different modulations can be found using the results for precoding loss in [51].

It is noted that the solution of processing matrices **B**, **F**, and **G** that satisfies (2.12) and (2.14) is not unique. In the conventional ZF nonlinear pre-processing algorithm proposed in [51] and [52], processing matrices **B**, **F**, and **G** are obtained using the following equations.

$$\mathbf{H}\mathbf{H}^H = \mathbf{R}\mathbf{R}^H \tag{2.15}$$

$$\mathbf{G}_C = \mathrm{diag}[[\mathbf{R}]_{11}^{-1}, [\mathbf{R}]_{22}^{-1}, \ldots, [\mathbf{R}]_{MM}^{-1}] \tag{2.16}$$

$$\mathbf{B}_C = \mathbf{G}\mathbf{R} \tag{2.17}$$

$$\mathbf{F}_C = \mathbf{H}^{\dagger}\mathbf{R} \tag{2.18}$$

where (2.15) is the Cholesky factorization [61] of $\mathbf{H}\mathbf{H}^H$, and **R** is an $M \times M$ lower triangular matrix with real diagonal elements.

Equivalently, matrices **B**, **F**, and **G** in the conventional ZF nonlinear pre-processing algorithm can also be obtained by performing the QR factorization [61] to $\mathbf{H}^H$, i.e.,

$$\mathbf{H}^H = \begin{bmatrix} \mathbf{Q} & \tilde{\mathbf{Q}} \end{bmatrix} \begin{bmatrix} \mathbf{R}' \\ \mathbf{0}_{(N-M),M} \end{bmatrix} \tag{2.19}$$

where $\mathbf{Q}$ is an $N \times M$ matrix, $\tilde{\mathbf{Q}}$ is an $N \times (N\text{-}M)$ matrix, and $\mathbf{R}'$ is an $M \times M$ upper triangular matrix with real diagonal elements. Then,

$$\mathbf{G}_C = \text{diag}[[\mathbf{R}']_{11}^{-1}, [\mathbf{R}']_{22}^{-1}, \ldots, [\mathbf{R}']_{MM}^{-1}] \tag{2.20}$$

$$\mathbf{B}_C = \mathbf{G}_C \mathbf{R}'^H \tag{2.21}$$

$$\mathbf{F}_C = \mathbf{Q}. \tag{2.22}$$

It can be easily found that the processing matrix $\mathbf{F}$ of the conventional ZF nonlinear pre-processing algorithm satisfies $\mathbf{F}_C^H \mathbf{F}_C = \mathbf{I}$, so the transmitted power constraint in (2.14) is satisfied.

From (2.11) and (2.12), one can see that if the ZF criterion is satisfied, $\mathbf{y}'$ is

$$\mathbf{y}' = \mathbf{v} + \tilde{\mathbf{n}}. \tag{2.23}$$

Therefore, all the inter-layer interference is eliminated. The performance of MS_i will be determined by the variance of the noise $\tilde{n}_i$. The covariance matrix of the noise vector $\tilde{\mathbf{n}}$ is

$$\begin{aligned} \mathbf{\Phi}_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} &\equiv \mathrm{E}[\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H] \\ &= \sigma_n^2 \mathbf{G}_C \mathbf{G}_C^H \\ &= \text{diag}[\sigma_n^2 / [\mathbf{R}]_{11}^2, \sigma_n^2 / [\mathbf{R}]_{22}^2, \ldots, \sigma_n^2 / [\mathbf{R}]_{MM}^2] \\ &= \text{diag}[\sigma_n^2 / [\mathbf{R}']_{11}^2, \sigma_n^2 / [\mathbf{R}']_{22}^2, \ldots, \sigma_n^2 / [\mathbf{R}']_{MM}^2] \end{aligned} \tag{2.24}$$

From (2.23) and (2.24), the SNR of MS_i's received signal y'_i is

$$\gamma_i = \mathrm{E}[|a_i|^2] / \mathrm{E}[|\tilde{n}_i|^2] = [\mathbf{R}]_{ii}^2 / \sigma_n^2. \tag{2.25}$$

Here, we assume the modulo device at the receiver can successfully remove the precoding vector $\mathbf{d}$ from the received effective data vector $\mathbf{v}$.

To examine the performance of different mobiles when the conventional ZF nonlinear pre-processing algorithm is applied, the following lemma from [71] and [72] is useful.

Lemma 2.1: Assume $\mathbf{A}$ is an $M \times N$ matrix ($M \leq N$) with i.i.d. entries, and each entry of $\mathbf{A}$ has a normalized zero-mean complex Gaussian distribution. $\mathbf{A} = \mathbf{LQ}$ is the QR-type decomposition where $\mathbf{L} = [l_{ij}]$ is an $M \times M$ lower triangular

matrix and $\mathbf{Q}$ is an $M \times N$ matrix with orthonormal rows. Then, the random variables $d_1 = l_{11}^2, d_2 = l_{22}^2, ..., d_M = l_{MM}^2$ are statistically independent and d_i has a central chi-square distribution with $2(M-i+1)$ degrees of freedom.

This lemma implies that for the conventional ZF nonlinear pre-processing algorithm described above the random variable $[\mathbf{R}]_{ii}^2$ has a central chi-square distribution with $2(M-i+1)$ degrees of freedom. Thus, the SNR of y'_i also has a central chi-square distribution with $2(M-i+1)$ degrees of freedom. As a result, the performance of MS_j is worse than that of MS_k if $j > k$.

The solution for processing matrices **B**, **F**, and **G** that satisfies (2.12) and (2.14) is not unique. However, in [51] and [52], there is no explanation about the reason why this particular solution for the processing matrices has been chosen. It is only mentioned in [52] that the processing matrices of the ZF DFD in the V-BLAST system [17], [18] are derived using the QR factorization, and the ZF DFD also applies a feedforward filter and a feedback filter to eliminate the inter-stream interference, so the authors of [51] and [52] propose to use the QR factorization to find the solution for the ZF-based nonlinear pre-processing algorithm. However, the feedforward filters in the nonlinear pre-processing structure of Figure 2.2 and in the ZF DFD structure have different functions due to the difference in their location. In the ZF DFD, the feedforward filter is at the receiver, so it influences the noise the DFD sees. The solution for the feedforward filter in the ZF DFD found using the QR factorization ensures that the noise is white and uncorrelated [64], [73]. However, in the nonlinear pre-processing structure, since the feedforward filter is located at the transmitter, the noise is not influenced by the feedforward filter, but the feedforward filter influences the transmitted power. Therefore, the design philosophy of the conventional ZF nonlinear pre-processing algorithm is somewhat questionable.

For the nonlinear pre-processing structure shown in Figure 2.2, the solution for the processing matrices that satisfies (2.12) and (2.14) is not unique, so it is important to know which solution should be chosen. In the next section, we

propose a generalized ZF nonlinear pre-pocessing structure, and investigate the optimal ZF nonlinear pre-processing solution.

2.4 Generalized ZF Nonlinear Pre-Processing Structure

The ZF criterion (2.12) for the nonlinear pre-processing structure shown in Figure 2.2 can be written equivalently as

$$\mathbf{HF} = \mathbf{G}^{-1}\mathbf{B}. \tag{2.26}$$

Recalling that **G** is a diagonal matrix, and **B** is a unit lower triangular matrix, from (2.26) it is clear that if matrix **F** is found, matrices **B** and **G** will result accordingly. Therefore, next, we will discuss how to design the matrix **F**.

First, it is clear that the ZF criterion (2.26) requires that the matrix **F** makes the product of **HF** a lower triangular matrix. Therefore, the kth column of **F** should be in the null space [61] of $\bar{\mathbf{H}}_k \equiv [(\mathbf{H})_1^T, (\mathbf{H})_2^T, ..., (\mathbf{H})_{k-1}^T]^T$, i.e.,

$$\bar{\mathbf{H}}_k [\mathbf{F}]_k = \mathbf{0}_{k-1,1}. \tag{2.27}$$

The solution of $[\mathbf{F}]_k$ that satisfies (2.27) can be represented as

$$[\mathbf{F}]_k = \mathbf{N}_k \mathbf{t}_k, \tag{2.28}$$

where $\mathbf{N}_k$ is an $N \times (N-k+1)$ matrix whose columns are the orthonormal basis vectors of null($\bar{\mathbf{H}}_k$) and $\mathbf{t}_k$ can be any $(N-k+1)\times 1$ vector. The matrix $\mathbf{N}_k$ can be found by singular value decomposition (SVD) [61] of $\bar{\mathbf{H}}_k$. Now, we can see that only $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ need to be determined to find **F**. It is noted that the matrix **F** should be designed under the transmitted power constraint in (2.14). Due to (2.28) and since $\mathbf{N}_k$ consists of orthonormal columns, then $\| [\mathbf{F}]_k \|^2 = \| \mathbf{t}_k \|^2$, and therefore (2.14) becomes

$$\sum_{k=1}^{M} \| \mathbf{t}_k \|^2 = M. \tag{2.29}$$

Consequently, $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ should be found under the constraint (2.29). Since there is an infinite number of sets $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ that satisfy (2.29), it is clear that there is an infinite number of solutions for the processing matrices that satisfy the ZF criterion under the transmitted power constraint (2.14). Now the problem of

how to choose $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ arises. To answer this question, we will first study how $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ influence the system performance.

From the ZF criterion (2.12), keeping in mind **B** is unit triangular, one can see that

$$[\mathbf{G}]_{kk}(\mathbf{H})_k[\mathbf{F}]_k = 1 \tag{2.30}$$

is required. Due to (2.28), (2.30) becomes

$$[\mathbf{G}]_{kk}(\mathbf{H})_k\mathbf{N}_k\mathbf{t}_k = 1. \tag{2.31}$$

From (2.23), (2.24), and (2.31), one can see that the SNR of MS_k's received signal is

$$\gamma_k = 1/(\sigma_n^2 |[\mathbf{G}]_{kk}|^2) = |(\mathbf{H})_k\mathbf{N}_k\mathbf{t}_k|^2/\sigma_n^2. \tag{2.32}$$

Therefore, the choice of vector $\mathbf{t}_k$ has direct influence on the performance of MS_k, and to optimize the performance of MS_k, $|(\mathbf{H})_k\mathbf{N}_k\mathbf{t}_k|^2$ should be maximized. However, the constraint (2.29) should be added when selecting $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$, and this constraint makes the optimization problems for different mobiles dependent.

In this thesis, we will solve the design problem of $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ in two steps. First, we give the optimal $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ solution under any given transmitted power allocation to different data streams. Then, the choice of the power allocation scheme will be studied.

2.4.1 The Generalized Structure

Since $\|\mathbf{t}_k\|^2 = \|[\mathbf{F}]_k\|^2$, one can see that the power of $\mathbf{t}_k$ is the transmitted power allocated to the *k*th transmitted data stream. Lemma 2.2 below gives the optimal $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ under any given transmitted power allocation to different data streams, and it can be easily proven.

Lemma 2.2: If the transmitted powers allocated to the *M* data streams are assumed to be $p_1, p_2, ..., p_M$ ($\|\mathbf{t}_k\|^2 = p_k$, $\sum_{k=1}^{M} p_k = M$), the vector $\mathbf{t}_k$ that minimizes the BER for MS_k is

$$\hat{\mathbf{t}}_k = \sqrt{p_k}\mathbf{N}_k^H(\mathbf{H})_k^H / \|(\mathbf{H})_k\mathbf{N}_k\|. \tag{2.33}$$

Using this result $[\mathbf{F}]_k$ becomes

$$[\mathbf{F}]_k = [\mathbf{F}_G]_k = \sqrt{p_k}\mathbf{N}_k\mathbf{N}_k^H(\mathbf{H})_k^H / \|(\mathbf{H})_k\mathbf{N}_k\|. \tag{2.34}$$

It can be shown that matrix $\mathbf{F}$ found using (2.34) can be represented equivalently as

$$\mathbf{F}_G = \mathbf{QP} \tag{2.35}$$

where $\mathbf{Q}$ is the QR factorization result of $\mathbf{H}^H$ in (2.19) and $\mathbf{P} = \text{diag}[\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_M}]$. Then, matrices $\mathbf{B}$ and $\mathbf{G}$ can be found to be

$$\mathbf{B}_G = \mathbf{P}^{-1}\mathbf{B}_C\mathbf{P} \tag{2.36}$$

$$\mathbf{G}_G = \mathbf{P}^{-1}\mathbf{G}_C \tag{2.37}$$

where $\mathbf{B}_C$ and $\mathbf{G}_C$ are matrices for the conventional ZF nonlinear pre-processing algorithm given by (2.21) and (2.20).

The feedforward filter $\mathbf{F}_G$ given by (2.35) is made up of two terms. The first term $\mathbf{Q}$ is determined by the channel matrix $\mathbf{H}$ (see (2.19)) and it is the same as the feedforward filter in the conventional ZF nonlinear pre-processing algorithm. The second term $\mathbf{P}$ is the power allocation matrix for the precoding output $\tilde{\mathbf{a}}$. Based on this solution, we propose a generalized ZF nonlinear pre-processing structure as shown in Figure 2.3. It is clear that different ZF nonlinear pre-processing algorithms can be realized by choosing different power allocation factors in this structure. From Lemma 2.2, one can see that the generalized ZF nonlinear pre-processing structure achieves minimum BER for each mobile under the power allocation $p_1, p_2, ..., p_M$.

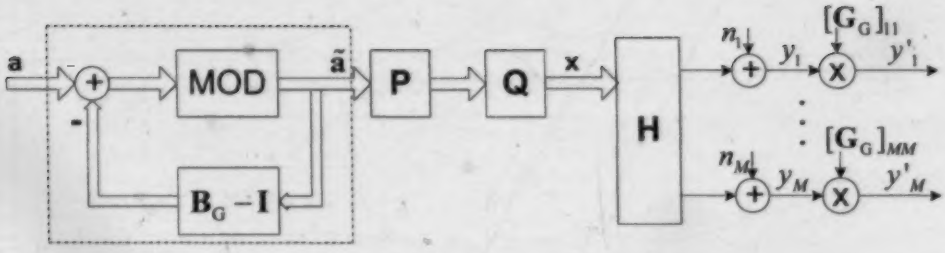


Figure 2.3 Structure of the generalized ZF nonlinear pre-processing.

The power allocation factors of the generalized ZF nonlinear pre-processing structure should be chosen based on different performance requirements. Next, we will discuss how to determine the power allocation factors.

2.4.2 Determining the Power Allocation Factors for the Generalized ZF Nonlinear Pre-Pocessing Sructure

First, it can be seen that the power allocation factors directly influence the performance of different mobiles. Using equations (2.35)–(2.37) for the processing matrices, the SNR of MS_k can be found as

$$\gamma_k = 1/(\sigma_n^2[\mathbf{G}_G]_{kk}^2) = p_k/(\sigma_n^2[\mathbf{G}_C]_{kk}^2). \tag{2.38}$$

2.4.2.1 Optimal ZF Nonlinear Pre-Processing Algorithm under any Relative SNR Requirement

From (2.38), it can be seen that the relative SNR performance of different mobiles can be controlled by power allocation, i.e., choosing the values of $p_1, p_2, ..., p_M$. For example, if the SNR of MS_k $(k = 1, 2, ..., M)$ is required to be $\overline{\gamma}_k = \alpha\lambda_k$, where λ_k relates to the ratio of different SNRs and α is a common factor, the power allocation factor should be set as

$$\overline{p}_k = [M/\sum_{k=1}^{M}(\lambda_k[\mathbf{G}_C]_{kk}^2)]\lambda_k[\mathbf{G}_C]_{kk}^2. \tag{2.39}$$

Interestingly, it can be further shown (see Appendix A) that the formulas (2.35) –(2.37) for the processing matrices with power allocation (2.39) achieve

minimum BER for each mobile for the required relative SNR defined by $\bar{\gamma}_k = \alpha\lambda_k$ for MS_k $(k = 1,2,...,M)$.

For a commercial wireless communication system where fair treatment of different users is desired, a ZF nonlinear pre-processing algorithm that can realize balanced performance of all mobiles is required. In this special case, the relative SNR of all the mobiles is required to be the same, i.e., $\bar{\gamma}_1 = \bar{\gamma}_2 = ... = \bar{\gamma}_M$. Using the above analysis, it can be seen that the processing matrices of the optimal ZF nonlinear pre-processing algorithm that achieves minimum BER for each mobile under the assumption that all the mobiles have equal SNR should be found using (2.35)–(2.37) with the following power allocation

$$p_k = [M / \sum_{k=1}^{M} ([\mathbf{G}_C]_{kk}^2)][\mathbf{G}_C]_{kk}^2 . \tag{2.40}$$

We will call the ZF nonlinear pre-processing algorithm using the processing matrices given by (2.35)–(2.37) and (2.40) the balanced ZF nonlinear pre-processing algorithm in this thesis.

2.4.2.2 Optimal ZF Nonlinear Pre-Processing Algorithm Satisfying the Individual SNR Target at Each Mobile

Another problem which is sometimes of interest in the wireless communications is to achieve minimum average total transmitted power when there is individual SNR target at each mobile, i.e., the SNR of MS_k is required to be larger than or equal to $\tilde{\gamma}_k$. It can be shown that the processing matrices (2.35)–(2.37) of the generalized ZF nonlinear pre-processing structure achieve the solution of this problem. The power allocation factor should be set to $\tilde{p}_k = \sigma_n^2 [\mathbf{G}_C]_{kk}^2 \tilde{\gamma}_k$, and the minimum average total transmitted energy per symbol interval then becomes

$$\tilde{\bar{E}}_{tr} = \sum_{k=1}^{M} \tilde{p}_k = \sigma_n^2 \sum_{k=1}^{M} ([\mathbf{G}_C]_{kk}^2 \tilde{\gamma}_k) . \tag{2.41}$$

2.4.2.3 Minimum BER ZF Nonlinear Pre-Processing Algorithm

From the average system performance point of view, the ZF-based nonlinear pre-processing algorithm that achieves minimum average BER (the minimum BER ZF nonlinear pre-processing algorithm) is of great interest. Since the generalized

ZF nonlinear pre-processing structure can achieve minimum average BER under any power allocation, it is clear that the minimum BER ZF nonlinear pre-processing algorithm can be realized by designing the power allocation matrix **P** of the generalized ZF nonlinear pre-processing structure. Next, we will derive the minimum BER ZF nonlinear pre-processing algorithm based on the generalized ZF nonlinear pre-processing structure we have proposed.

Assume the BER of MS_k is $P(\gamma_k)$. From (2.38) the average BER of all mobiles is

$$P_e = (1/M)\sum_{k=1}^{M} P(\gamma_k) = (1/M)\sum_{k=1}^{M} P(p_k/(\sigma_n^2[\mathbf{G}_C]_{kk}^2)). \qquad (2.42)$$

The problem is to find the values of $\hat{p}_1, \hat{p}_2, ..., \hat{p}_M$ that minimize P_e. It is seen that the problem of finding $\hat{p}_1, \hat{p}_2, ..., \hat{p}_M$ that minimize P_e has the same form as the problem of finding the power allocation factors that minimize the average BER for the ZF DFD in the V-BLAST system [75]. Therefore, the method proposed in [75] can be used to solve the problem for the minimum BER ZF nonlinear pre-processing algorithm. For BPSK and QPSK with Gray encoding [74], exact expression of the average BER can be found and then the average BER is minimized. For modulations of larger sizes, it is hard to obtain the exact expression of the average BER, so an approximate expression for the average BER is minimized [75].

2.4.3 Relationship between the Generalized ZF Nonlinear Pre-Processing Structure and the Conventional ZF Nonlinear Pre-Processing Algorithm

Comparing the solutions for the processing matrices of the conventional ZF nonlinear pre-processing algorithm and the generalized ZF nonlinear pre-processing structure, it can be seen clearly that the conventional ZF nonlinear pre-processing algorithm is a special implementation of the generalized ZF nonlinear pre-processing structure with $\mathbf{P} = \mathbf{I}$, which means the transmitted power is equally allocated to the M data streams. Therefore, from Lemma 2.2, one can see that the conventional ZF nonlinear pre-processing algorithm minimizes the BER for each

mobile under equal power allocation. To our knowledge, this property of the conventional ZF nonlinear pre-processing algorithm has not been previously demonstrated.

It is noted that the processing matrices of the conventional ZF nonlinear pre-processing algorithm are intuitively introduced. The influence of this algorithm on the performance of different mobiles and on the overall performance of the system has not been considered during the design. From the analysis in Section 2.3 one can see that using this algorithm the SNRs of different mobiles are decided by the channel matrix. It can be found that there are significant differences in the performance of different mobiles, and the gap between the worst performance and the best performance of the mobiles is very large. Also, this algorithm is not optimal from the average system performance point of view. The generalized ZF nonlinear pre-processing structure we propose provides a framework to design different ZF nonlinear pre-processing algorithms based on different optimization requirements, including optimization under any given relative SNR requirement, optimization under the individual SNR target, and optimization in the average system BER performance sense, as shown in Sections 2.4.2.1–2.4.2.3.

It is also interesting to point out that the proposed generalized ZF nonlinear pre-processing structure provides a simplified way to achieve power allocation for the conventional ZF nonlinear pre-processing algorithm. In [76], a nonlinear pre-processing structure that supports power allocation is proposed for single-user MIMO systems. This structure can be directly used to perform power allocation in the conventional ZF nonlinear pre-processing algorithm. With this structure, power allocation works on the original data vector **a**, so data symbols with different powers are input into the modulo device. The modulo device therefore has to be modified to support different signal amplitudes. Assume the powers of different data symbols are assigned as $\hat{p}_1, \hat{p}_2, ..., \hat{p}_M$ ($\sum_{k=1}^{M} \hat{p}_k = M$). From (2.25), the SNR of MS_k is then $\gamma_k = \hat{p}_k / (\sigma_n^2 [\mathbf{G}_C]_{kk}^2)$. Comparing this result with (2.38), it is clear that the generalized ZF nonlinear pre-processing structure can realize the

same SNRs by choosing $\mathbf{P} = \text{diag}[\sqrt{\hat{p}_1}, \sqrt{\hat{p}_2}, ..., \sqrt{\hat{p}_M}]$. It can also be proven that the same maximum achievable sum rate as in the conventional ZF nonlinear pre-processing algorithm using the structure of [76] can be achieved using the proposed structure. Since power allocation works on $\tilde{\mathbf{a}}$ instead of $\mathbf{a}$ in the proposed structure, the modulo device needs no change, and hence, the system complexity is reduced.

2.4.4 Simulation Results

In this section, we use simulation to show the performance advantage of the balanced ZF nonlinear pre-processing algorithm proposed in Section 2.4.2.1 in improving the worst mobile's performance compared to the conventional ZF nonlinear pre-processing algorithm, and the performance advantage of the minimum BER ZF nonlinear pre-processing algorithm proposed in Section 2.4.2.3 in improving the average BER performance of the system compared to the conventional ZF nonlinear pre-processing algorithm.

In our simulations, BER vs. SNR curves are determined. We assume the channel matrix $\mathbf{H}$ changes independently from one symbol interval to the next. The SNR used in our simulation is defined as

$$\text{SNR} \equiv \frac{\overline{E}_{tr}}{N_0 \log_2 M_c} \tag{2.43}$$

where $\overline{E}_{tr} = \text{E}[\| \mathbf{x} \|^2]$ is the average total transmitted energy per symbol interval, N_0 is the one-sided noise power spectral density, and M_c is the constellation size. The transmitted power increase of the TH precoder is accounted for with this SNR definition.

2.4.4.1 Performance Comparison of the Balanced ZF Nonlinear Pre-Processing Algorithm and the Conventional ZF Nonlinear Pre-Processing Algorithm

From the analysis in Section 2.3, it can be found that when the conventional ZF nonlinear pre-processing algorithm is applied, the performance of MS_j is worse than that of MS_k if $j > k$. It can be shown by simulation that the gap between the worst performance and the best performance of the mobiles is very large. This is

an undesirable feature for a commercial wireless communication system where fair treatment of different users is usually required. The balanced ZF nonlinear pre-processing algorithm proposed in Section 2.4.2.1 can solve this problem. We have shown that this algorithm can achieve minimum BER for each mobile under the assumption that all the mobiles have equal SNR. Next, we will use simulation to show the advantage of the balanced ZF nonlinear pre-processing algorithm. The number of single-antenna mobiles and the number of transmit antennas at the base station are chosen as $M = N = 4$. BPSK modulation is used.

Figure 2.4 shows the performance comparison of the balanced ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm. When the conventional ZF nonlinear pre-processing algorithm is used, the BER curves of $MS_1,\ldots,MS_4$ are obtained. It can be found that MS_4, whose signal component is precoded latest, has the worst performance and the performance gaps between different mobiles are very large. One can see that at BER = 10^{-3}, the performance of MS_4 is approximately 21 dB worse than the performance of MS_1. When the balanced ZF nonlinear pre-processing algorithm is used, all the mobiles have the same performance and it can be clearly seen that the worst mobile's performance is greatly improved. At BER = 10^{-3}, the balanced ZF nonlinear pre-processing algorithm achieves approximately 5.5 dB gain over the worst mobile's performance. In Figure 2.4, the average mobile performance of the conventional ZF nonlinear pre-processing algorithm and the balanced ZF nonlinear pre-processing algorithm are also compared. It is noted that when the balanced ZF nonlinear pre-processing algorithm is used, all the mobiles have the same performance, so the average mobile performance is the same as the performance of each mobile. It can be found that although the average mobile performance of the balanced ZF nonlinear pre-processing algorithm exhibits a little degradation as compared to that of the conventional ZF nonlinear pre-processing algorithm, the two algorithms' average mobile performance is very close at high SNR. Considering the remarkable ability of the balanced ZF nonlinear pre-processing algorithm to improve the worst mobile's performance, it is a very promising algorithm.

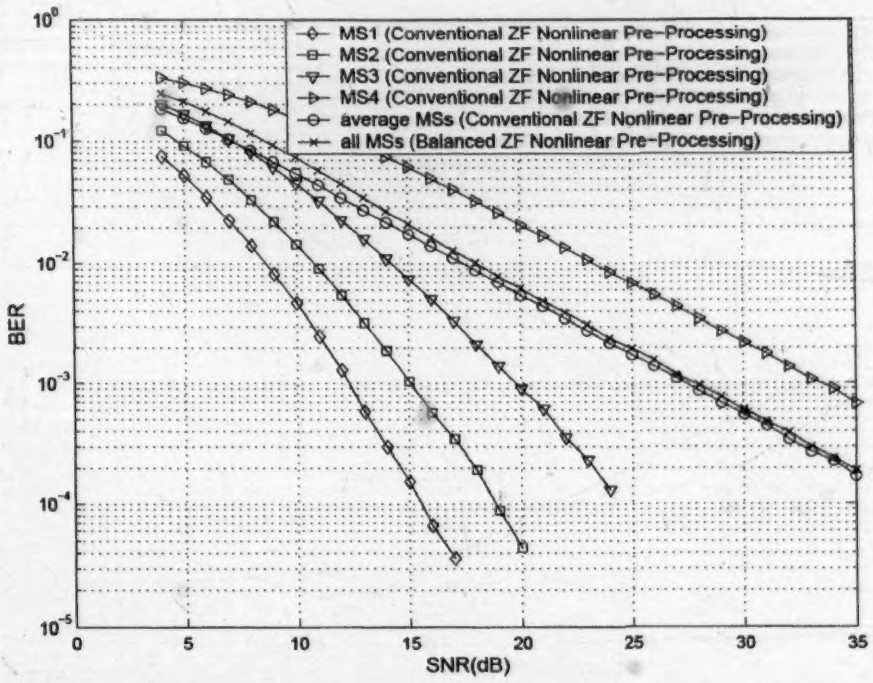


Figure 2.4 Performance comparison of the balanced ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (BPSK, $M = N = 4$).

2.4.4.2 Performance Comparison of the Minimum BER ZF Nonlinear Pre-Processing Algorithm and the Conventional ZF Nonlinear Pre-Processing Algorithm

The advantage of the minimum BER ZF nonlinear pre-processing algorithm over the conventional ZF nonlinear pre-processing algorithm is demonstrated by simulation in Figure 2.5 and Figure 2.6. In Figure 2.5, a multi-user MIMO system with 4 transmit antennas at the base station and 4 single-antenna mobiles ($M = N = 4$) is considered, while in Figure 2.6 $M = N = 8$. The average BER over all the mobiles vs. SNR curves are determined with BPSK modulation. It is seen that the minimum BER ZF nonlinear pre-processing algorithm achieves significantly better performance than the conventional ZF nonlinear pre-processing algorithm. At BER = 10^{-3}, the minimum BER ZF nonlinear pre-processing algorithm gives

about 5 dB gain over the conventional ZF nonlinear pre-processing algorithm for the system with $M = N = 4$. For the system with $M = N = 8$, the minimum BER ZF nonlinear pre-processing algorithm gives about 7.5 dB gain over the conventional ZF nonlinear pre-processing algorithm at BER = 10^{-3}.

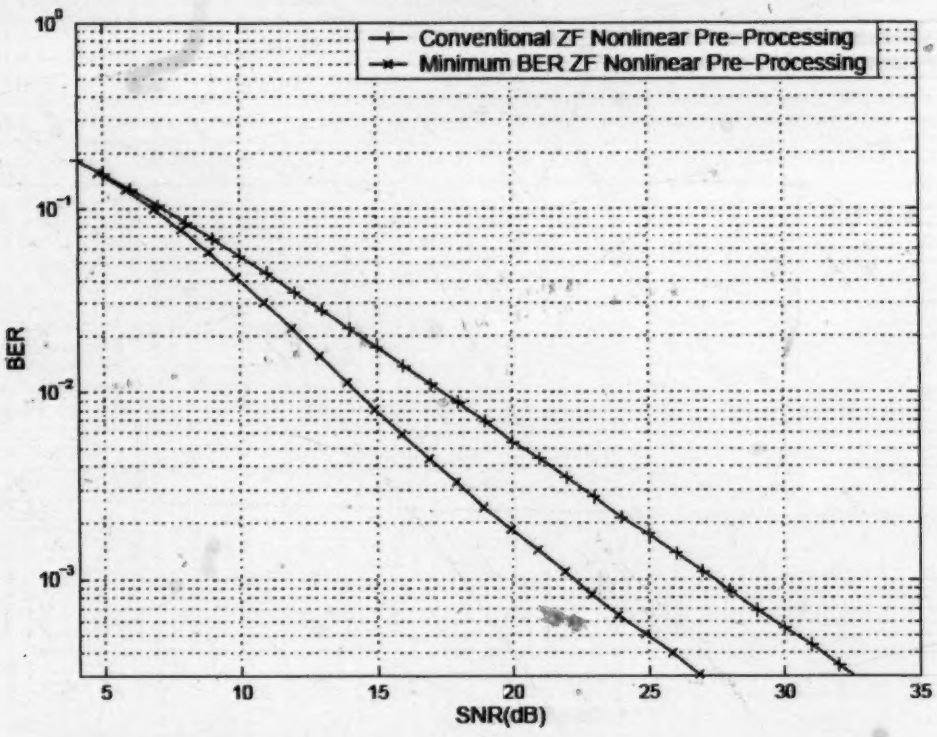


Figure 2.5 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (BPSK, $M = N = 4$).

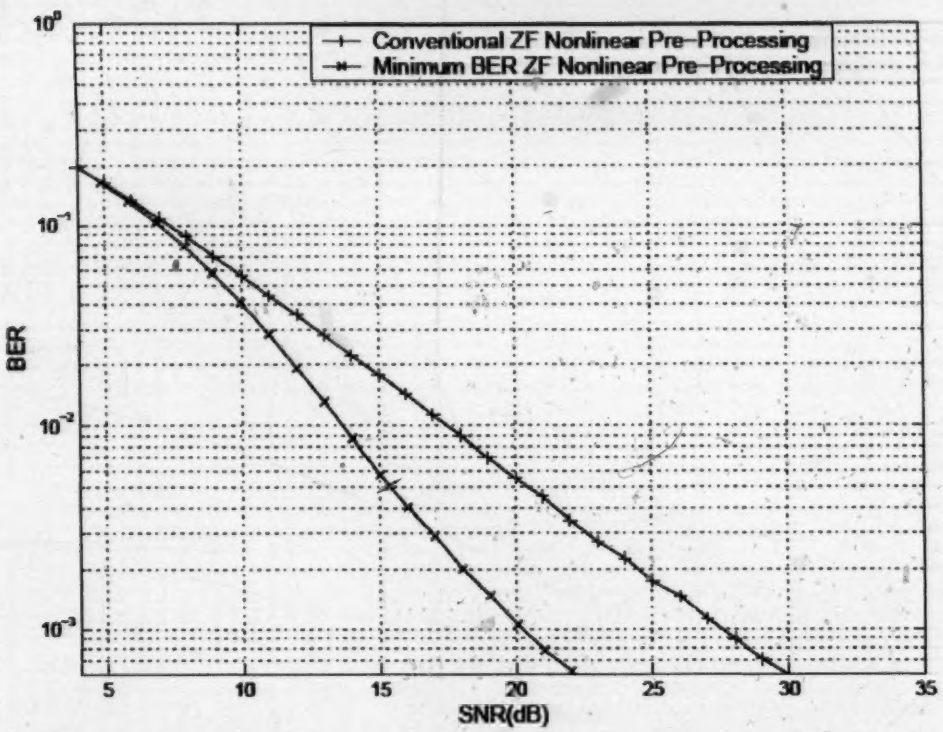


Figure 2.6 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (BPSK, $M = N = 8$).

2.5 MMSE Nonlinear Pre-Processing Algorithm

For the nonlinear pre-processing structure shown in Figure 2.2, the application of the ZF criterion to design the processing matrices has been discussed in Section 2.4. However, although the ZF criterion (2.12) ensures that all the inter-stream interference can be pre-eliminated, it does not constrain the noise enhancement. To achieve better performance, in this section, the MMSE criterion is used to design the processing matrices of the nonlinear pre-processing structure in Figure 2.2. Leaving some residual inter-stream interference, the MMSE criterion based algorithm (MMSE nonlinear pre-processing algorithm) minimizes the MSE, which includes both the residual inter-stream interference and noise. Thus, it can

achieve significantly better performance than the ZF-based nonlinear pre-processing algorithm.

Next, we will derive the MMSE nonlinear pre-processing algorithm for the system structure shown in Figure 2.2. It is assumed that the transmitter has the knowledge of the noise variance σ_n^2.

The error vector that needs to be considered for the system in Figure 2.2 is the difference between the effective data vector $\mathbf{v} = \mathbf{a} + \mathbf{d}$ and the signal vector $\mathbf{y}'$ entering the decision module , i.e.,

$$\mathbf{e} = \mathbf{y}' - \mathbf{v} = \mathbf{G}\mathbf{y} - \mathbf{B}\tilde{\mathbf{a}} = (\mathbf{GHF} - \mathbf{B})\tilde{\mathbf{a}} + \mathbf{Gn} \tag{2.44}$$

The error covariance matrix is

$$\begin{aligned} \mathbf{\Phi}_{ee} &\equiv \mathrm{E}[\mathbf{e}\mathbf{e}^H] \\ &= (\mathbf{GHF} - \mathbf{B})\mathbf{R}_{\tilde{\mathbf{a}}}(\mathbf{GHF} - \mathbf{B})^H + \sigma_n^2 \mathbf{G}\mathbf{G}^H \end{aligned} \tag{2.45}$$

The MMSE solution should be able to minimize $\mathrm{tr}(\mathbf{\Phi}_{ee})$ with respect to processing matrices **B**, **F**, and **G**, where **B** is a unit lower triangular matrix and **G** is a diagonal matrix.

First, as has been revealed in Section 2.3, the pre-processing matrix **F** will influence the transmitted power. We still use the average transmitted power constraint, so matrix **F** is required to satisfy (2.14), i.e., $\mathrm{tr}\{\mathbf{F}\mathbf{F}^H\} = M$. To satisfy the transmitted power constraint (2.14) and to obtain a closed-from expression for the processing matrices, we assume that matrix **F** has orthonormal columns, i.e., $\mathbf{F}^H\mathbf{F} = \mathbf{I}$. It is noted that the processing matrix **F** in the conventional ZF nonlinear pre-processing algorithm described in Section 2.3 also has orthonormal columns, and with this assumption the transmitted power is equally allocated to the multiple data streams to be transmitted. As it will be discussed in Section 2.5.1, the benefit of this assumption is that it can avoid the variation of the transmitted power at each transmit antenna for systems with M=N.

Invoking the orthogonality principle, the MMSE solution should satisfy

$$\mathrm{E}[\mathbf{e}\mathbf{y}^H] = 0. \tag{2.46}$$

Thus, from (2.44)

$$\mathrm{E}[\mathbf{G}\mathbf{y}\mathbf{y}^H - \mathbf{B}\tilde{\mathbf{a}}\mathbf{y}^H] = 0 \tag{2.47}$$

$$\mathbf{G}(\mathbf{H}\mathbf{F}\mathbf{R}_{\tilde{a}}\mathbf{F}^H\mathbf{H}^H + \sigma_n^2\mathbf{I}) = \mathbf{B}\mathbf{R}_{\tilde{a}}\mathbf{F}^H\mathbf{H}^H. \quad (2.48)$$

Introducing $\xi = \sigma_n^2 / \sigma_a^2$, (2.48) becomes

$$\mathbf{G}(\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H + \xi\mathbf{I}) = \mathbf{B}\mathbf{F}^H\mathbf{H}^H. \quad (2.49)$$

Our task now is to find matrices **B**, **F**, and **G** that can satisfy (2.49), where **B** is an $M \times M$ unit lower triangular matrix, **F** is an $N \times M$ matrix with orthonormal columns, and **G** is an $M \times M$ diagonal matrix.

It is found that matrices **B**, **F**, and **G** that satisfy (2.49) can be derived using the following equations.

$$\mathbf{R}\mathbf{R}^H = (\mathbf{H} + \xi\mathbf{H}^{\dagger H})(\mathbf{H}^H + \xi\mathbf{H}^{\dagger}) \quad (2.50)$$

$$\mathbf{G} = \mathrm{diag}[[\mathbf{R}]_{11}^{-1}, [\mathbf{R}]_{22}^{-1}, \ldots, [\mathbf{R}]_{MM}^{-1}] \quad (2.51)$$

$$\mathbf{B} = \mathbf{G}\mathbf{R} \quad (2.52)$$

$$\mathbf{F} = (\mathbf{H}^H + \xi\mathbf{H}^{\dagger})\mathbf{R}^{-H}. \quad (2.53)$$

Here, (2.50) is the Cholesky factorization of $(\mathbf{H} + \xi\mathbf{H}^{\dagger H})(\mathbf{H}^H + \xi\mathbf{H}^{\dagger})$, and **R** is an $M \times M$ lower triangular matrix.

Equivalently, the above result of **B**, **F**, and **G** can also be obtained through the QR factorization of $\mathbf{H}^H + \xi\mathbf{H}^{\dagger}$. Assuming the QR factorization of $\mathbf{H}^H + \xi\mathbf{H}^{\dagger}$ gives

$$\mathbf{H}^H + \xi\mathbf{H}^{\dagger} = \begin{bmatrix}\mathbf{Q}' & \tilde{\mathbf{Q}}'\end{bmatrix}\begin{bmatrix}\mathbf{R}' \\ \mathbf{0}_{(N-M),M}\end{bmatrix} \quad (2.54)$$

where $\mathbf{Q}'$ is an $N \times M$ matrix, $\tilde{\mathbf{Q}}'$ is an $N \times (N\text{-}M)$ matrix, and $\mathbf{R}'$ is an $M \times M$ upper triangular matrix, matrices **B**, **F**, and **G** are

$$\mathbf{G} = \mathrm{diag}[[\mathbf{R}']_{11}^{-1}, [\mathbf{R}']_{22}^{-1}, \ldots, [\mathbf{R}']_{MM}^{-1}] \quad (2.55)$$

$$\mathbf{B} = \mathbf{G}\mathbf{R}'^H \quad (2.56)$$

$$\mathbf{F} = \mathbf{Q}'. \quad (2.57)$$

Proof that matrices **B**, **F**, and **G** given by (2.50)–(2.53) or (2.54)–(2.57) can satisfy (2.49) is included in Appendix B.

With the **B**, **F**, and **G** we found above for the MMSE nonlinear pre-processing algorithm, the error covariance matrix is

$$\Phi_{ee} = \sigma_n^2 \mathbf{G}^2 + \sigma_n^2 \xi \mathbf{G}(\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{G} \tag{2.58}$$

Appendix C shows how (2.58) is derived.

Finally, let us compare the computational complexity of the proposed MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm of [51] and [52]. From Section 2.3, it can be seen that the computational complexity of the conventional ZF nonlinear pre-processing algorithm is mainly determined by the QR factorization of an $N \times M$ matrix, $\mathbf{H}^H$. From the derivation of the MMSE nonlinear pre-processing algorithm in this section, it is seen that the computational complexity of the MMSE nonlinear pre-processing algorithm is mainly determined by the calculation of the pseudo-inverse of matrix $\mathbf{H}$ and the QR factorization of an $N \times M$ matrix, $\mathbf{H}^H + \xi \mathbf{H}^\dagger$. Therefore, although the computational complexity of the proposed MMSE nonlinear pre-processing algorithm is higher than that of the conventional ZF nonlinear pre-processing algorithm, the computational complexity of the former is still of the same order as that of the latter, e.g., in systems with $M = N$, the computational complexity of both algorithms is $O(N^3)$.

In Section 2.5.2, we will use simulations to show that the MMSE nonlinear pre-processing algorithm we have developed in this section can achieve significant performance improvement over the conventional ZF nonlinear pre-processing algorithm.

2.5.1 Comparison of the Proposed MMSE Nonlinear Pre-Processing Algorithm and Other MMSE-Criterion-Based Design

The design of the MMSE-criterion-based nonlinear pre-processing algorithm which minimizes the sum MSE at all the receive branches under a fixed total transmitted power has also been studied by some other recent work [66], [77]–[80]. The MMSE nonlinear pre-processing algorithm proposed in this thesis is different from the work in [66], [77]–[80] and exhibits its unique advantage.

For simplicity we will call the algorithm of [77]–[79] as the JBU MMSE nonlinear pre-processing algorithm[2] in the sequel. The difference between the proposed MMSE nonlinear pre-processing algorithm and the JBU MMSE nonlinear pre-processing algorithm is discussed in the following.

The MMSE nonlinear pre-processing algorithm we have proposed is derived under the constraint that the processing matrix **F** has orthonormal columns. With this constraint, when $M = N$, **F** becomes a unitary matrix, so the transmitted power at each transmit antenna is determined only by $\sigma_{\tilde{a}}^2$, i.e.,

$$\mathrm{E}[|x_n|^2] = \mathrm{E}[|(\mathbf{F})_n \tilde{\mathbf{a}}|^2] = \sigma_{\tilde{a}}^2 \qquad n = 1,...,N. \tag{2.59}$$

Since $\sigma_{\tilde{a}}^2 = \gamma_p^2 \sigma_a^2$ and the value of γ_p^2 is already known [51], in this case the required dynamic range of the transmitted power at each antenna is not influenced by the pre-processing. The transmitted powers at the multiple transmit antennas will all have the same value and this value will not change with the variation of the channel. This feature is an important advantage in the design of a practical system. If $\mathrm{E}[|x_n|^2]$ varies, the linear dynamic range of the power amplifier at each antenna must be able to cover the variation of $\mathrm{E}[|x_n|^2]$. Otherwise, the output signal will be distorted. Because at high SNR the sum capacity of the multi-user MIMO downlink grows linearly with min(M,N) [23]–[26] the systems with $M = N$ are most efficient from the capacity's point of view. Therefore, this feature of the MMSE nonlinear pre-processing algorithm is very useful.

The expressions for the processing matrices in the JBU MMSE nonlinear pre-processing algorithm of [77]–[79] are derived based on the assumption that all the scaling factors at different mobiles have the same value; this constraint is not used in the MMSE nonlinear pre-processing we propose. It is noted that this constraint is not necessary for the precoding algorithm to work. Its adoption in [77]–[79] is to facilitate the derivation of a closed-form expression for processing matrices. The processing matrix **F** in the JBU MMSE nonlinear pre-processing algorithm is required to make the total transmitted power constant, while no other constraint is

---

[2] JBU indicates the names of the authors of [77], where the algorithm was originally introduced.

added on **F**. Therefore, the MMSE nonlinear pre-processing algorithm we have proposed and the JBU MMSE nonlinear pre-processing algorithm in fact deal with two optimization problems under different constraints, which leads to different derivation and different solutions for the processing matrices **B**, **F**, and **G**. The processing matrix **F** of the JBU MMSE nonlinear pre-processing algorithm does not have orthonormal columns. Therefore, when the JBU MMSE nonlinear pre-processing algorithm is applied in systems with $M = N$ the transmitted power at each transmit antenna is different and varies with the variation of the channel. It will be shown by simulation in Section 2.5.2 that the transmitted power at each transmit antenna spans a very large range. Therefore, when this algorithm is applied, the required dynamic range of the power amplifier at each antenna is much higher than that needed with the MMSE nonlinear pre-processing algorithm we have proposed.

The MMSE-criterion-based nonlinear pre-processing algorithm proposed in [66] is derived based on the duality of multi-user MIMO downlink with dirty-paper coding and multi-user MIMO uplink with perfect SIC. The processing matrices of this algorithm have to be found by exhaustive search. More recently, processing matrices for the nonlinear pre-processing based on the MMSE criterion have been solved by exploiting the duality between DFD and THP in [80]. However, an iterative method needs to be used to find the solution, and as stated in [80] "the global optimality of the solution" cannot be proved. Since the processing matrices of the MMSE nonlinear pre-processing algorithm we have proposed are defined by closed-form expressions, the computational complexity of the proposed algorithm is significantly lower than that of the algorithms of [66] and [80].

2.5.2 Simulation Results

In our simulations, BER vs. SNR curves are determined. The BER is the average BER over all the mobiles. Two kinds of modulation, QPSK and square 16QAM with Gray encoding [74], are used. The channel is assumed to change independently from one symbol interval to the next. The SNR used in our simulation is defined by (2.43).

Figure 2.7 shows the performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm when QPSK with Gray encoding [74] is used. The number of single-antenna mobiles and the number of transmit antennas at the base station are chosen as $M = N = 4$. It is clearly seen that the MMSE nonlinear pre-processing algorithm can achieve significantly better performance than the conventional ZF nonlinear pre-processing algorithm. At BER = 10^{-3} the MMSE nonlinear pre-processing algorithm achieves approximately 7 dB gain over the conventional ZF nonlinear pre-processing algorithm.

Figure 2.8 presents the performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm when 16QAM with Gray encoding [74] is used. The number of single-antenna mobiles and the number of transmit antennas at the base station are $M = N = 4$. It is seen that at BER = 10^{-3} the MMSE nonlinear pre-processing algorithm gives about 3.5 dB gain over the conventional ZF nonlinear pre-processing algorithm.

In Figure 2.9, the performance curves of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm in a multi-user MIMO system with 6 transmit antennas at the base station and 4 single-antenna mobiles are drawn. QPSK with Gray encoding [74] is used. It can be observed that at BER = 10^{-3} the MMSE nonlinear pre-processing algorithm gives approximately 1.8 dB gain over the conventional ZF nonlinear pre-processing algorithm.

It is well known that the MMSE filter converges to the ZF filter when SNR goes to infinity or noise variance goes to zero [74]. However, the performance of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm does not converge in the high SNR region as shown in Figure 2.7, Figure 2.8, and Figure 2.9. It has also been shown in [30] and [44] that when working in the MIMO antenna systems the performance of the MMSE linear detector does not converge to that of the ZF linear detector at high SNR region, and the performance of the MMSE DFD also does not converge to that of the ZF DFD at high SNR region. This somewhat puzzling phenomenon has been

studied in [81] and [82]. It has been shown there that the gap between the post-processing SNRs of the ZF DFD and the MMSE DFD when working in MIMO antenna systems does not diminish as SNR goes to infinity, which implies the performance gap between these two algorithms does not vanish at high SNR region. The MMSE DFD can gain more post-processing signal power than the ZF DFD, because instead of nulling out all the inter-stream interference as the ZF filter, the MMSE filter can effectively recover the desired signal's information hidden in the interference space. Using similar approach as that of [81] and [82], one can find that the performance gap between the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm also does not diminish in high SNR region.

Finally, the performance of the MMSE nonlinear pre-processing algorithm and the JBU MMSE nonlinear pre-processing algorithm of [77]–[79] is compared. A multi-user MIMO system with 4 transmit antennas at the base station and 4 single-antenna mobiles is considered, and 16QAM is used. Based on the discussion in Section 2.5.1, we know that when the MMSE nonlinear pre-processing algorithm is applied the transmitted power at each antenna will not change with the variation of the channel, but when the JBU MMSE nonlinear pre-processing algorithm is applied the transmitted power at each antenna will vary with variations of the channel. Simulation is used to show how large is the dynamic range of the transmitted power at each antenna when the JBU MMSE nonlinear pre-processing algorithm is applied. 10^6 channel realizations have been tested. The dynamic range of the transmitted power at each antenna, which is denoted by D, is shown in Table 2.1 for different SNRs. It can be seen that the dynamic range of the transmitted power at each antenna is quite large when the JBU MMSE nonlinear pre-processing algorithm is applied. For example, when the SNR is 29.2 dB, our simulation shows that the dynamic range of the transmitted power at each antenna is as high as 27.6 dB, which implies the maximum transmitted power is 575 times the minimum transmitted power at each transmit antenna. The BER performance of the two algorithms is compared in Figure 2.10. The JBU MMSE nonlinear pre-processing algorithm shows better

performance than the MMSE nonlinear pre-processing algorithm we have proposed. However, it should be noted that the JBU MMSE nonlinear pre-processing algorithm has the drawback of requiring large dynamic range of the transmitted power at each antenna. Therefore, the MMSE nonlinear pre-processing algorithm we have proposed appears to have a greater potential in practical applications.

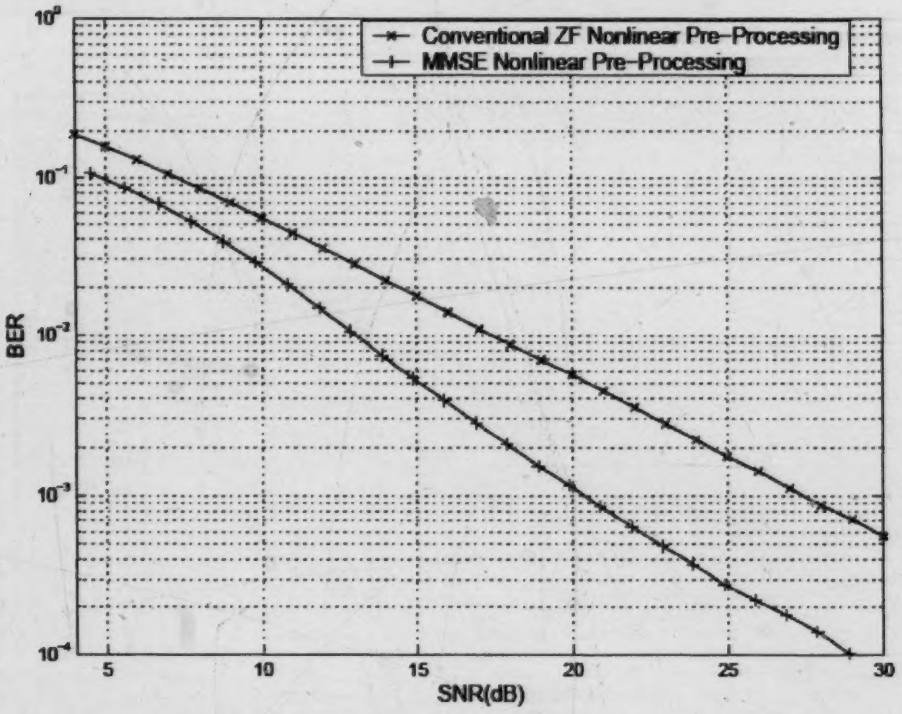


Figure 2.7 Performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $M = N = 4$).

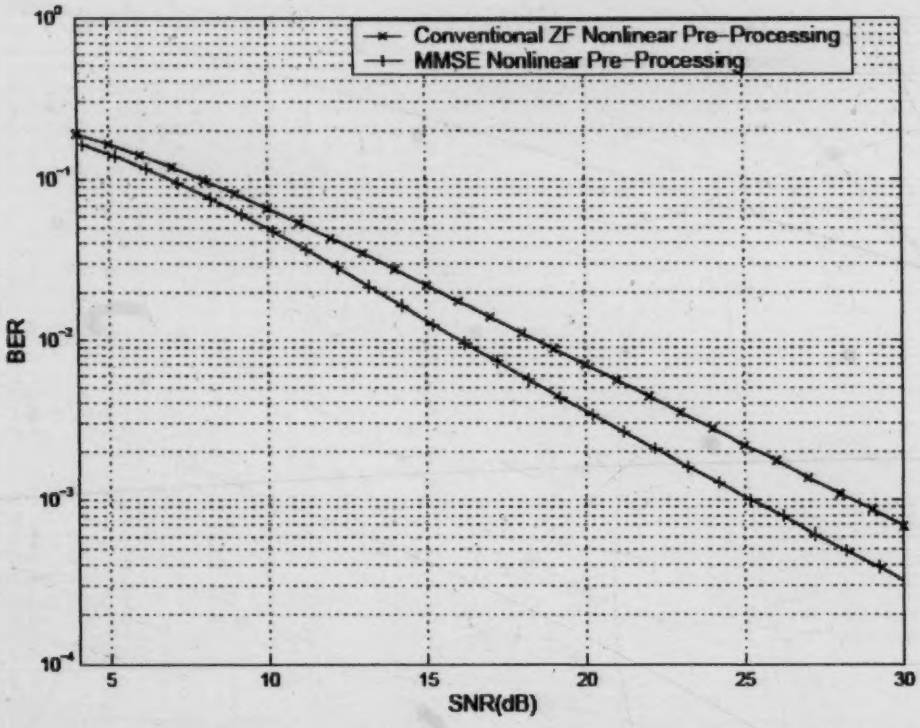


Figure 2.8 Performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (16QAM, $M = N = 4$).

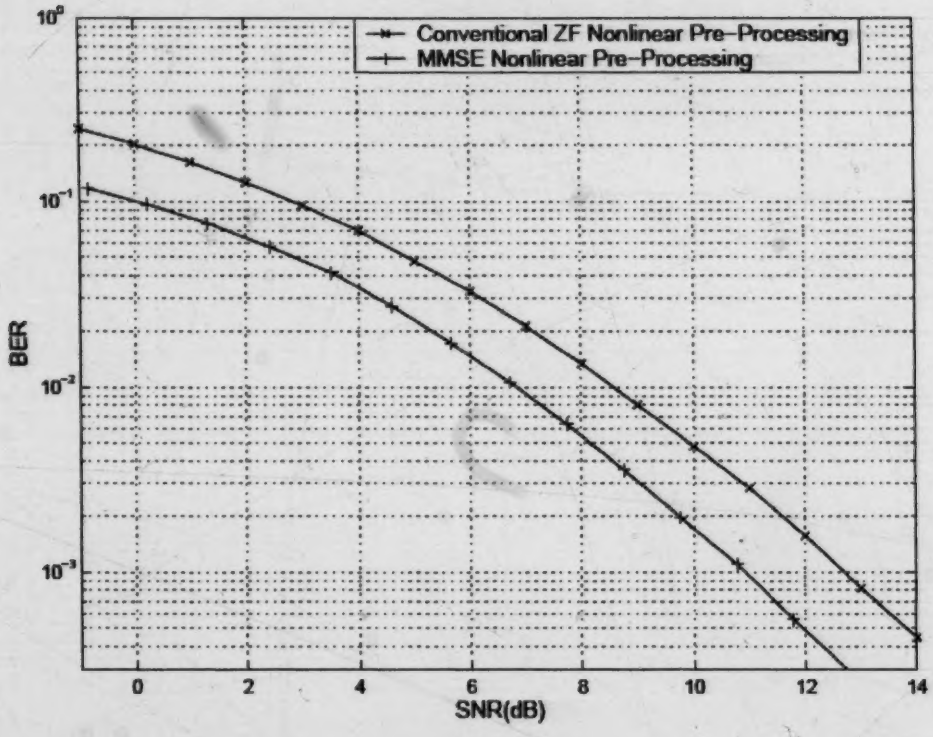


Figure 2.9 Performance comparison of the MMSE nonlinear pre-processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $M = 4$, and $N = 6$).

Table 2.1 The dynamic range of the transmitted power at each antenna when the JBU MMSE nonlinear pre-processing algorithm is applied.

SNR(dB)	4.2	5.2	6.2	7.2	8.2	9.2	10.2	11.2	12.2
D(dB)	22.3	17.7	18.5	17.4	20.9	18.1	16.1	18.7	19.9
SNR(dB)	13.2	14.2	15.2	16.2	17.2	18.2	19.2	20.2	21.2
D(dB)	17.1	17.9	20.5	19.1	18.3	19.6	19.9	19.9	22.3
SNR(dB)	22.2	23.2	24.2	25.2	26.2	27.2	28.2	29.2	30.2
D(dB)	22.0	22.5	23.9	23.4	24.7	25.2	26.2	27.6	27.5

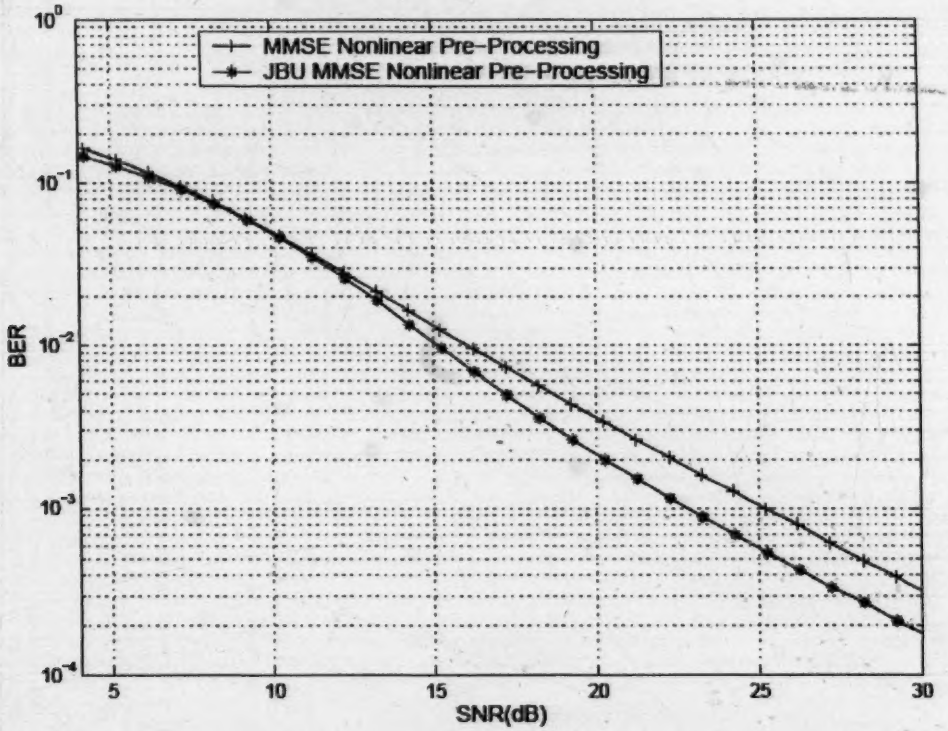


Figure 2.10 Performance comparison of the proposed MMSE nonlinear pre-processing algorithm and the JBU MMSE nonlinear pre-processing algorithm (16QAM, $M = N = 4$).

Chapter 3 Optimal Ordering Lemma

In this chapter, we will first show the importance of ordering in the nonlinear pre-processing algorithms described and proposed in the last chapter. The change of the system structure caused by ordering will be revealed. Then, an optimal ordering lemma will be introduced. Using this lemma, we will show that the "best-first" ordering method can achieve the optimal order in the minimax noise variance sense for the conventional ZF nonlinear pre-processing algorithm and the optimal order in the minimax error variance sense for the MMSE nonlinear pre-processing algorithm we propose. An efficient way to perform the "best-first" ordering for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm will be presented. The ordering problem in the other ZF-based nonlinear pre-processing algorithms proposed in Section 2.4 will also be addressed.

The system structure and assumptions described in Section 2.1 will be used in this chapter.

3.1 The Importance of Ordering

When the conventional ZF nonlinear pre-processing algorithm described in the last chapter is applied, it has been found that the performance of the system is decided by the noise variances $\sigma_{\tilde{n}_1}^2,...,\sigma_{\tilde{n}_M}^2$. From (2.19)–(2.22) and (2.24), one can see that the values of $\sigma_{\tilde{n}_1}^2,...,\sigma_{\tilde{n}_M}^2$ are decided by the QR decomposition of the conjugate transpose of the channel matrix, $\mathbf{H}^H$. It can be easily observed that ordering the columns of channel matrix $\mathbf{H}$ will not change the noise variances

$\sigma_{\tilde{n}_1}^2,...,\sigma_{\tilde{n}_M}^2$. Thus, the system performance will not change due to ordering the columns of **H**. However, ordering the rows of **H** will lead to corresponding change of $\sigma_{\tilde{n}_1}^2,...,\sigma_{\tilde{n}_M}^2$. Thus, the system performance will change accordingly. For example, let us consider a 4 × 4 channel matrix

$$\mathbf{H}=\begin{bmatrix} 0.03-0.91i & 0.82-0.51i & -0.83+0.30i & 0.51+0.32i \\ 0.26-0.89i & -1.84+0.68i & 0.56-0.43i & 0.29-0.53i \\ -0.21+0.43i & 0.19+0.16i & 0.34+1.44i & 0.32-0.22i \\ -0.21+0.00i & 0.34+0.12i & 0.70+0.32i & -1.04-0.87i \end{bmatrix} \tag{3.1}$$

Using (2.19)–(2.22) and (2.24), one can find that the noise variances $\sigma_{\tilde{n}_1}^2,\sigma_{\tilde{n}_2}^2,\sigma_{\tilde{n}_3}^2,\sigma_{\tilde{n}_4}^2$ are

$$(\sigma_{\tilde{n}_1}^2,\sigma_{\tilde{n}_2}^2,\sigma_{\tilde{n}_3}^2,\sigma_{\tilde{n}_4}^2)=\sigma_n^2(0.34,0.22,0.62,1.01). \tag{3.2}$$

If the rows of **H** are ordered into

$$\mathbf{H}'=[(\mathbf{H})_4^T,(\mathbf{H})_3^T,(\mathbf{H})_1^T,(\mathbf{H})_2^T]^T$$

$$=\begin{bmatrix} -0.21+0.00i & 0.34+0.12i & 0.70+0.32i & -1.04-0.87i \\ -0.21+0.43i & 0.19+0.16i & 0.34+1.44i & 0.32-0.22i \\ 0.03-0.91i & 0.82-0.51i & -0.83+0.30i & 0.51+0.32i \\ 0.26-0.89i & -1.84+0.68i & 0.56-0.43i & 0.29-0.53i \end{bmatrix} \tag{3.3}$$

using **H'** instead of **H** in (2.19)–(2.22), one can get

$$(\sigma_{\tilde{n}_4}^2,\sigma_{\tilde{n}_3}^2,\sigma_{\tilde{n}_1}^2,\sigma_{\tilde{n}_2}^2)=\sigma_n^2(0.38,0.57,0.50,0.43) \tag{3.4}$$

Therefore, it is clear that the performance of the conventional ZF nonlinear pre-processing algorithm might change if the order of the rows of the channel matrix changes.

This phenomenon can be explained intuitively as follows. For the conventional ZF nonlinear pre-processing algorithm, the pre-processing can be seen as an iterative procedure. For the first precoded signal, the inter-stream interference from all the other signals is nulled out by linear precoding. Then, for the ith (i = 2, ..., M-1) precoded signal, the spatial interference from the (i+1)th, ..., Mth precoded signals is nulled out by linear precoding, and the TH precoder is used to pre-eliminate the spatial interference from the 1st, ..., (i-1)th precoded signals. For the last precoded signal, the spatial interference from all the

other signals is pre-eliminated by the TH precoder. Nulling out the spatial interference from other users using linear precoding will lead to corresponding noise enhancement at the receiver. Hence, choosing different groups of signals to be nulled out will lead to different results of the noise enhancement. Therefore, different precoding orders of the signals will result in different noise enhancement effects and thus different performance. For the nonlinear pre-processing structure shown in Figure 2.2, it can be seen that ordering the rows of the channel matrix will directly change the precoding order. Therefore, the rows of the channel matrix can be ordered to minimize the negative impact of the noise enhancement on the system performance, or in other words to improve the system performance.

The same phenomenon can also be observed for the MMSE nonlinear pre-processing algorithm and the other ZF nonlinear pre-processing algorithms based on the generalized ZF nonlinear pre-processing structure we have proposed in Sections 2.5 and 2.4, i.e., ordering the columns of the channel matrix will not change the system performance, but the system performance might change if the rows of the channel matrix are ordered in a different way.

Therefore, if the rows of the channel matrix are suitably ordered, better performance can be achieved for the conventional ZF nonlinear pre-processing algorithm, the MMSE nonlinear pre-processing algorithm, and the generalized ZF nonlinear pre-processing structure based algorithms. In this chapter, we intend to find ordering methods on the rows of $\mathbf{H}$ which can lead to performance improvement for these algorithms.

3.2 Change of the System Structure Due to Ordering

First, let us study how the ordering of the rows of $\mathbf{H}$ will influence the system structure.

The rows of channel matrix $\mathbf{H} \equiv [(\mathbf{H})_1^T, (\mathbf{H})_2^T, ..., (\mathbf{H})_M^T]^T$ are indexed by $1,2,\ldots,M$ where $(\mathbf{H})_i$ represents the channel gains from the Ant_1, Ant_2,…,Ant_N to MS_i. We use $L \equiv \{l_1, l_2, ..., l_M\}$ to represent an order of the rows of $\mathbf{H}$, and it is one permutation of the row indices, $1,2,\ldots,M$. It should be noted that in order to use the optimal ordering lemma introduced later, $l_1, l_2, ..., l_M$ denote the Mth row, the

$(M\text{-}1)$th row, …, the 1st row of the ordered $\mathbf{H}$, respectively, i.e., if the order of the rows of $\mathbf{H}$ is $L = \{l_1, l_2, ..., l_M\}$, $\mathbf{H}$ should be ordered into

$$\mathbf{H}^{(L)} = [(\mathbf{H})_{l_M}^T, (\mathbf{H})_{l_{M-1}}^T, ..., (\mathbf{H})_{l_1}^T]^T. \tag{3.5}$$

We assume that the data symbol a_i is destined for the MS_i. For the nonlinear pre-processing structure shown in Figure 2.2, if no ordering is applied, $\mathbf{a} = (a_1, a_2, ..., a_M)^T$ is the data vector at the input of the precoding device. Since processing matrix $\mathbf{B}$ is a lower triangular matrix, precoding is processed from a_1, a_2 to a_M. $\mathbf{x} = (x_1, x_2, ..., x_N)^T = \mathbf{F}\mathbf{B}^{-1}(\mathbf{a} + \mathbf{d})$ is the signal vector to be transmitted by the N transmit antennas. This structure is shown in Figure 3.1.

If $\mathbf{H}$ is ordered into $\mathbf{H}^{(L)}$ as in (3.5), processing matrices $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{G}$ should be generated using $\mathbf{H}^{(L)}$ instead of $\mathbf{H}$, and they are denoted as $\mathbf{B}^{(L)}$, $\mathbf{F}^{(L)}$, and $\mathbf{G}^{(L)}$. Now, the received signal vector is represented as $\mathbf{y}^{(L)} = \mathbf{H}^{(L)}\mathbf{x} + \mathbf{n}^{(L)}$, where $\mathbf{y}^{(L)} = (y_{l_M}, y_{l_{M-1}}, ..., y_{l_1})^T$, $\mathbf{n}^{(L)} = (n_{l_M}, n_{l_{M-1}}, ..., n_{l_1})^T$. In this situation, to transmit a_i to MS_i, $\mathbf{a}^{(L)}$ should be $\mathbf{a}^{(L)} = (a_{l_M}, a_{l_{M-1}}, ..., a_{l_1})^T$, so precoding is processed from a_{l_M}, $a_{l_{M-1}}$ to a_{l_1} as shown in Figure 3.2. $\mathbf{x} = (x_1, x_2, ..., x_N)^T = \mathbf{F}^{(L)}\mathbf{B}^{(L)-1}(\mathbf{a}^{(L)} + \mathbf{d}^{(L)})$ is the signal vector to be transmitted by the N transmit antennas.

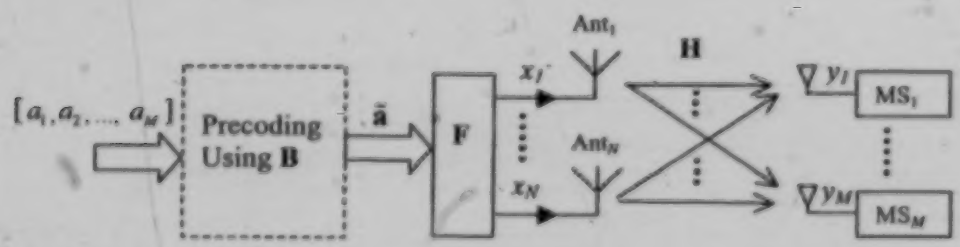


Figure 3.1 Nonlinear pre-processing structure for the downlink of a multi-user MIMO system without ordering.

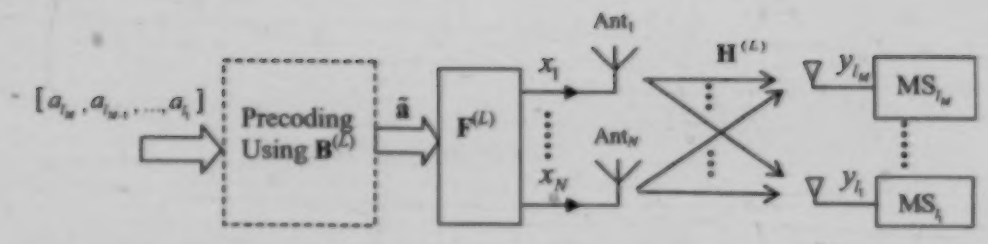


Figure 3.2 Nonlinear pre-processing structure for the downlink of a multi-user MIMO system if order L is used.

3.3 Optimal Ordering Lemma

A "best-first" ordering method has been used in the ZF DFD and MMSE DFD for single-user MIMO systems in [18] and [28]. A proof of the "best-first" ordering achieving the optimal order in the minimax noise variance sense for the ZF DFD was given in [27]. However, a proof of the optimum of the "best-first" ordering in the MMSE DFD was not given. In the following, an optimal ordering lemma will be proposed. This lemma gives the conditions under which the "best-first" ordering method is optimal. We have used this lemma to show that the "best-first" ordering achieves the optimal order in the minimax error variance sense for the MMSE DFD [44]. In this chapter, we will use this lemma to show that the "best-first" ordering method can achieve the optimal order in the minimax noise variance sense for the conventional ZF nonlinear pre-processing algorithm and the optimal order in the minimax error variance sense for the MMSE nonlinear pre-processing algorithm we have proposed in Section 2.5.

To introduce the optimal ordering lemma, first let us assume there is a set of components indexed by 1,2,...,M that need to be ordered. We define the *performance parameters* $p_1, p_2, ..., p_M$ as the parameters that relate to the M components and directly reflect the performance of the order. The values of these performance parameters vary with the change of the order of these components, so better performance can be achieved by ordering these components. For example, for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm the components that need to be

ordered are the rows of the channel matrix **H**. The performance parameters are the noise variances ($E[|\tilde{n}_1|^2]$, $E[|\tilde{n}_2|^2]$,..., $E[|\tilde{n}_M|^2]$) for the conventional ZF nonlinear pre-processing algorithm, and the error variances

$$(\varepsilon_1, \varepsilon_2, ..., \varepsilon_M) \equiv (E[|e_1|^2], E[|e_2|^2], ..., E[|e_M|^2]) = \text{diag}(\Phi_{ee}) \qquad (3.6)$$

for the MMSE nonlinear pre-processing algorithm. The optimal ordering lemma is as follows.

Optimal Ordering Lemma: For the components indexed by 1,2,…,M that need to be ordered, if the performance parameters of these components, $p_1, p_2, ..., p_M$, obey the *Condition* 1 and *Condition 2* described below, the order $L \equiv \{l_1, l_2, ..., l_M\}$ found by the "best-first" ordering is the optimal "best-worst" order with regard to the performance parameters. This means that the worst performance parameter p_{l_i} when order L is used is better than or the same as the worst performance parameter p_{q_i} when any other order $Q \equiv \{q_1, q_2, ..., q_M\}$ is used.

"Best" and "worst" in the above lemma should be decided according to particular application. For example, if the components need to be ordered are rows of the channel matrix **H** for the conventional ZF nonlinear pre-processing algorithm, the performance parameters are the noise variances ($E[|\tilde{n}_1|^2]$, $E[|\tilde{n}_2|^2]$,..., $E[|\tilde{n}_M|^2]$), so "best" means minimum and "worst" means maximum. In this case, the optimal "best-worst" order with regard to the performance parameters means the optimal order in the minimax noise variance sense.

The two conditions of the optimal ordering lemma are:

Condition 1: For any two orders, $A \equiv \{A_1, A_2, ..., A_M\}$ and $B \equiv \{B_1, B_2, ..., B_M\}$, of the indices of the M components to be ordered, if $A_k = B_k$ and their *constraint sets* $\hat{C}_{A_k}$ and $\hat{C}_{B_k}$ are made up of the same elements (regardless of the order of these elements, e.g., {2,3,1} and {2,1,3} are made up of the same elements), then $p_{A_k} = p_{B_k}$.

Condition 2: For any two orders, $A \equiv \{A_1, A_2, ..., A_M\}$ and $B \equiv \{B_1, B_2, ..., B_M\}$, if $A_k = B_p$ and the constraint set $\hat{C}_{A_k} \subseteq \hat{C}_{B_p}$, then p_{A_k} is better than or the same as p_{B_p}.

The definition of the constraint set is: For an order $L \equiv \{l_1, l_2, ..., l_M\}$, the constraint set $\hat{C}_{l_i}$ of l_i is $\hat{C}_{l_i} = \{l_i, l_{i+1}, ..., l_M\}$. A similar definition can be found in [18].

For the components indexed by 1,2,...,M to be ordered, the "best-first" ordering finds the final order $L \equiv \{l_1, l_2, ..., l_M\}$ from l_1, l_2, to l_M iteratively. To find l_i, all the unordered components (the components whose indices have not been chosen as $l_1, ..., l_{i-1}$) are tested. The index of the component with the best performance parameter if it is chosen as l_i will be ordered into the first available position being ordered, i.e., the position l_i. For this reason it is called the "best-first" ordering method. The procedure of its operation is described as follows:

Initialization:

1. $i = 1$,

2. Create a set: $\hat{U} = \{1, 2, ..., M\}$.

Iteration until $i > M$:

1. From all the components whose indices belong to $\hat{U}$, find the component that has the best p_{l_i} if it is put in the position i of the order L,

2. Put the index of the component found in 1 (assume it is α) in position i of the order L, i.e., $l_i = \alpha$,

3. Delete α from $\hat{U}$,

4. $i = i+1$.

Appendix D gives the proof of this lemma.

3.4 Ordering in the Conventional ZF Nonlinear Pre-Processing Algorithm and the MMSE Nonlinear Pre-Processing Algorithm

For the conventional ZF nonlinear pre-processing algorithm, it has been revealed in Section 2.3 that the performance of MS_i is determined by the variance of the noise $\tilde{n}_i$. Therefore, the system performance will be dominated by the largest noise variance, and hence we intend to find the optimal order in the minimax noise variance sense. That means, if $L \equiv \{l_1, l_2, ..., l_M\}$ is the optimal order, the maximal noise variance in $(E[|\tilde{n}_{l_1}|^2], E[|\tilde{n}_{l_2}|^2], ..., E[|\tilde{n}_{l_M}|^2])$ when order L is used must be smaller than or the same as the maximal noise variance when any other order is used. Similarly, for the MMSE nonlinear pre-processing algorithm, the error variances are considered, and our goal is to achieve the optimal order in the minimax error variance sense. That means, if $L \equiv \{l_1, l_2, ..., l_M\}$ is the optimal order, the maximal error variance in $(\varepsilon_{l_1}, \varepsilon_{l_2}, ..., \varepsilon_{l_M})$ when order L is used must be smaller than or the same as the maximal error variance when any other order is used.

For the conventional ZF nonlinear pre-processing algorithm, the noise variances $(E[|\tilde{n}_{l_1}|^2], E[|\tilde{n}_{l_2}|^2], ..., E[|\tilde{n}_{l_M}|^2])$ are used as the performance parameters in the ordering. In the same manner, for the MMSE nonlinear pre-processing algorithm, the error variances $(\varepsilon_{l_1}, \varepsilon_{l_2}, ..., \varepsilon_{l_M})$ are used as the performance parameters. It can be shown that for both the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm, the performance parameters obey the following two conditions, i.e.,

Condition 1: For any two orders of the rows of $\mathbf{H}$, $A \equiv \{A_1, A_2, ..., A_M\}$ and $B \equiv \{B_1, B_2, ..., B_M\}$, if $A_k = B_k$ and the constraint sets $\hat{C}_{A_k}$ and $\hat{C}_{B_k}$ are made by same elements, then $E[|\tilde{n}_{A_k}|^2] = E[|\tilde{n}_{B_k}|^2]$ or $\varepsilon_{A_k} = \varepsilon_{B_k}$.

Condition 2: For any two orders of the rows of **H**, $A \equiv \{A_1, A_2, ..., A_M\}$ and $B \equiv \{B_1, B_2, ..., B_M\}$, if $A_k = B_p$ and the constraint sets $\hat{C}_{A_k} \subseteq \hat{C}_{B_p}$, then $E[|\tilde{n}_{A_k}|^2] \le E[|\tilde{n}_{B_p}|^2]$ or $\varepsilon_{A_k} \le \varepsilon_{B_p}$.

Employing the optimal ordering lemma, we see that the "best-first" ordering can achieve the optimal order in the minimax noise variance sense for the conventional ZF nonlinear pre-processing algorithm and the optimal order in the minimax error variance sense for the MMSE nonlinear pre-processing algorithm. (Here, "best" means minimum and "worst" means maximum, so the "best-worst" order is the minimax order.)

3.4.1 An Efficient Way to Perform the "Best-First" Ordering for the Conventional ZF Nonlinear Pre-Processing Algorithm and the MMSE Nonlinear Pre-Processing Algorithm

From the description of the "best-first" ordering in Section 3.3, one can see that the optimal order $L \equiv \{l_1, l_2, ..., l_M\}$ is found from l_1, l_2 to l_M. To determine the ith element of the optimal order, the performance parameter p_{l_i}, when each unordered component is put in the position l_i, needs to be calculated. In other words, assuming $U = \{u_1, u_2, ..., u_P\}$ is made up of the indices of the components that have not been ordered yet ($P = M-i+1$), the performance parameters when $l_i = u_1, u_2, ..., u_P$ need to be found. We use $p_{l_i(u_j)}$ to represent the performance parameter if $l_i = u_j$ ($j = 1,...,P$). Since performance parameters vary with the change of the order, generally, we need to form the order with $l_i = u_j$ and then calculate $p_{l_i(u_j)}$ for $j = 1,...,P$, respectively. In [28], an efficient method has been proposed to calculate $p_{l_i(u_j)}$ ($j = 1,...,P$) when the "best-first" ordering is applied to the MMSE DFD. Using this method, instead of forming different orders by choosing $l_i = u_j$ and then doing the calculation, one can find all the values of $p_{l_i(u_j)}$ ($j = 1,...,P$) from one equation. Thus, the computational complexity is greatly reduced.

As we have explained, the noise variances are used as the performance parameters in the "best-first" ordering for the conventional ZF nonlinear pre-processing algorithm and the error variances are used as the performance parameters in the "best-first" ordering for the MMSE nonlinear pre-processing algorithm. During the "best-first" ordering of these two algorithms, to determine l_i, the value of noise variance $E[|\tilde{n}_{l_i}|^2]$ or error variance ε_{l_i}, when each unordered row of $\mathbf{H}$ is put in the position l_i, needs to be calculated. We use $E[|\tilde{n}_{l_i(u_j)}|^2]$ to represent the noise variance if $l_i = u_j$ ($j = 1,\ldots,P$) for the conventional ZF nonlinear pre-processing algorithm, and $\varepsilon_{l_i(u_j)}$ to represent the error variance if $l_i = u_j$ ($j = 1,\ldots,P$) for the MMSE nonlinear pre-processing algorithm. We find that as in the "best-first" ordering for the MMSE DFD in [28], the calculation of $E[|\tilde{n}_{l_i(u_j)}|^2]$ or $\varepsilon_{l_i(u_j)}$ ($j = 1,\ldots,P$) can also be simplified. The following equations are proposed to calculate $E[|\tilde{n}_{l_i(u_j)}|^2]$ or $\varepsilon_{l_i(u_j)}$ ($j = 1,\ldots,P$).

First, we form

$$\mathbf{H}_U \equiv [(\mathbf{H})_{u_1}^T, (\mathbf{H})_{u_2}^T, \ldots, (\mathbf{H})_{u_P}^T]^T \tag{3.7}$$

$$\mathbf{D}_U \equiv [[\mathbf{H}^\dagger]_{u_1}, [\mathbf{H}^\dagger]_{u_2}, \ldots, [\mathbf{H}^\dagger]_{u_P}] \tag{3.8}$$

where $[\mathbf{H}^\dagger]_k$ represents the kth column of $\mathbf{H}^\dagger$.

For the conventional ZF nonlinear pre-processing algorithm, the values of $E[|\tilde{n}_{l_i(u_1)}|^2]$, $E[|\tilde{n}_{l_i(u_2)}|^2]$,..., $E[|\tilde{n}_{l_i(u_P)}|^2]$ are

$$(E[|\tilde{n}_{l_i(u_1)}|^2], E[|\tilde{n}_{l_i(u_2)}|^2], \ldots, E[|\tilde{n}_{l_i(u_P)}|^2]) = \sigma_n^2 \mathrm{diag}((\mathbf{H}_U \mathbf{H}_U^H)^{-1}) \tag{3.9}$$

For the MMSE nonlinear pre-processing algorithm, the value of $\varepsilon_{l_i(u_j)}$ is

$$\varepsilon_{l_i(u_j)} = \sigma_n^2 [\mathbf{T}]_{jj} + \sigma_n^2 \xi [\mathbf{T}]_{jj} [\mathbf{S}]_{u_j u_j} \tag{3.10}$$

where

$$\mathbf{T} = [(\mathbf{H}_U + \xi \mathbf{D}_U^H)(\mathbf{H}_U^H + \xi \mathbf{D}_U)]^{-1} \tag{3.11}$$

$$\mathbf{S} = (\mathbf{H}\mathbf{H}^H)^{-1} \tag{3.12}$$

Appendix E shows how (3.9) and (3.10)–(3.12) are derived. Using the above equations, the computational complexity of performing the "best-first" ordering for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm is significantly reduced.

3.4.2 Simulation Results

In this section, simulations are used to show the advantage of the "best-first" ordering in the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm. BER vs. SNR curves are determined. The BER is the average BER over all the mobiles. The number of single-antenna mobiles and the number of transmit antennas at the base station are chosen as $M = N = 4$. The channel is assumed to change independently from one symbol interval to the next. The SNR used in our simulation is defined by (2.43).

Figure 3.3 shows the performance comparison of the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is applied. QPSK with Gray encoding [74] is used. It is clearly seen that the "best-first" ordering can significantly improve the system performance for both algorithms. At BER = 10^{-3}, using the "best-first" ordering, approximately 5.5 dB and 3.5 dB gain is obtained for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm, respectively.

Figure 3.4 shows the performance improvement achieved by the "best-first" ordering for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm when 16QAM with Gray encoding [74] is used. It is seen that the "best-first" ordering achieves about 5 dB and 6 dB gain for the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm, respectively, at BER = 10^{-3}.

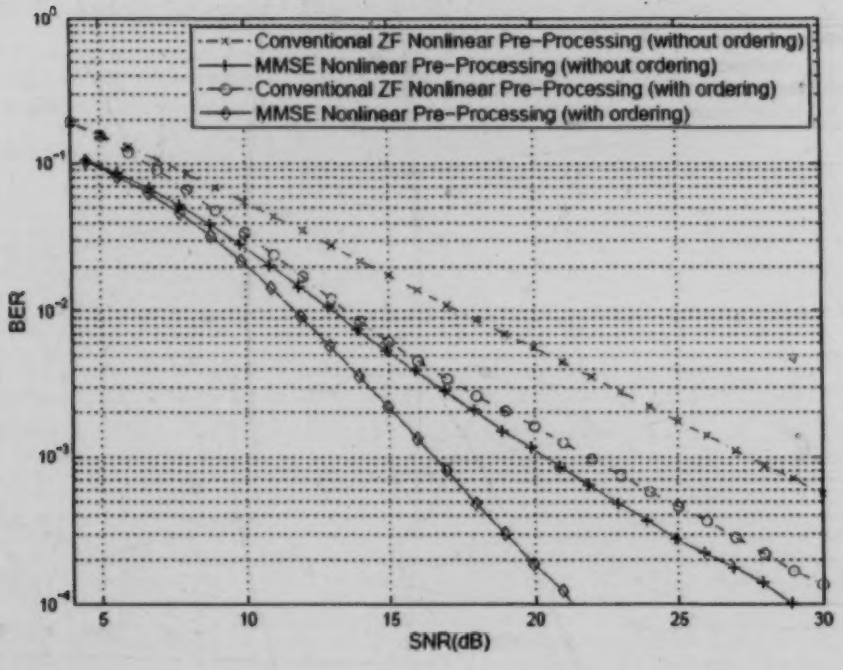


Figure 3.3 Performance comparison of the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is applied (QPSK, $M = N = 4$).

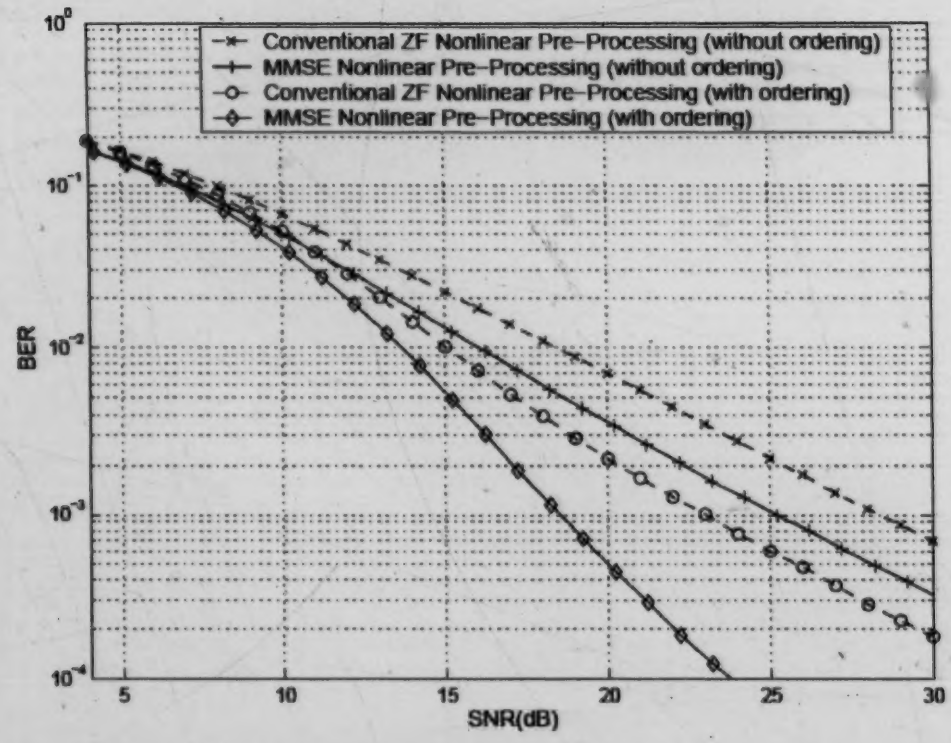


Figure 3.4 Performance comparison of the conventional ZF nonlinear pre-processing algorithm and the MMSE nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is applied (16QAM, $M = N = 4$).

3.5 Ordering in Other ZF-Based Nonlinear Pre-Processing Algorithms

In Section 2.4, a generalized ZF nonlinear pre-processing structure has been proposed. This structure can realize ZF-based nonlinear pre-processing algorithm that achieves minimum BER for each mobile under any given power allocation to different data streams. Based on the generalized ZF nonlinear pre-processing structure, different ZF-based nonlinear pre-processing algorithms have been developed. These algorithms include the optimal ZF nonlinear pre-processing algorithm under the relative SNR requirement, the balanced ZF nonlinear pre-processing algorithm, the optimal ZF nonlinear pre-processing algorithm

satisfying the individual SNR target at each mobile, and the minimum BER ZF nonlinear pre-processing algorithm. It can be easily found that the performance of these algorithms can also be improved if the rows of the channel matrix are ordered properly, but the purposes of ordering are different in these algorithms due to their different optimization requirements.

For the optimal ZF nonlinear pre-processing algorithm under the relative SNR requirement proposed in Section 2.4.2.1, assume the relative SNR requirement is defined by $\bar{\gamma}_k = \alpha\lambda_k$ for MS_k $(k = 1,2,...,M)$. If no ordering is applied, one can find that the relative SNR at MS_k is

$$\bar{\gamma}_k = [M/(\sigma_n^2 \sum_{k=1}^{M} (\lambda_k [\mathbf{G}_C]_{kk}^2))]\lambda_k \tag{3.13}$$

where $\mathbf{G}_C$ is found from (2.15)–(2.18). If the rows of the channel matrix are ordered into $\mathbf{H}^{(L)} \equiv [(\mathbf{H})_{l_M}^T, (\mathbf{H})_{l_{M-1}}^T, ..., (\mathbf{H})_{l_1}^T]^T$ and by using $\mathbf{H}^{(L)}$ the matrix $\mathbf{G}_C$ becomes $\mathbf{G}_C^{(L)}$, the relative SNR at MS_k is

$$\bar{\gamma}_k = [M/(\sigma_n^2 \sum_{k=1}^{M} (\lambda_k [\mathbf{G}_C^{(L)}]_{kk}^2))]\lambda_k . \tag{3.14}$$

Clearly, in this case the optimal order should be the one that can minimize the $\sum_{k=1}^{M} (\lambda_k [\mathbf{G}_C^{(L)}]_{kk}^2)$.

The balanced ZF nonlinear pre-processing algorithm is a special case of the optimal ZF nonlinear pre-processing algorithm under the relative SNR requirement with $\bar{\gamma}_1 = \bar{\gamma}_2 = ... = \bar{\gamma}_M$. Therefore, the optimal order should be the one that can minimize the $\sum_{k=1}^{M} ([\mathbf{G}_C^{(L)}]_{kk}^2)$ where $\mathbf{G}_C^{(L)}$ is found from (2.15)–(2.18) by using the ordered channel matrix $\mathbf{H}^{(L)}$.

For the optimal ZF nonlinear pre-processing algorithm satisfying the individual SNR target $\tilde{\gamma}_k$ at each mobile proposed in Section 2.4.2.2, if no ordering is applied, one can find that the minimum average total transmitted energy per symbol interval is

$$\tilde{\bar{E}}_{tr} = \sum_{k=1}^{M} \tilde{p}_k = \sigma_n^2 \sum_{k=1}^{M} ([\mathbf{G}_C]_{kk}^2 \tilde{\gamma}_k). \quad (3.15)$$

If the rows of the channel matrix are ordered into $\mathbf{H}^{(L)}$ and by using $\mathbf{H}^{(L)}$ the matrix $\mathbf{G}_C$ becomes $\mathbf{G}_C^{(L)}$, the minimum average total transmitted energy per symbol interval is

$$\tilde{\bar{E}}_{tr} = \sum_{k=1}^{M} \tilde{p}_k = \sigma_n^2 \sum_{k=1}^{M} ([\mathbf{G}_C^{(L)}]_{kk}^2 \tilde{\gamma}_k). \quad (3.16)$$

Clearly, in this case the optimal order should be the one that can minimize the $\sum_{k=1}^{M} ([\mathbf{G}_C^{(L)}]_{kk}^2 \tilde{\gamma}_k)$.

For the minimum BER ZF nonlinear pre-processing algorithm proposed in Section 2.4.2.3, if no ordering is applied, one can find that the minimum average BER of all mobiles is

$$P_e = (1/M) \sum_{k=1}^{M} P(\hat{p}_k / (\sigma_n^2 [\mathbf{G}_C]_{kk}^2)) \quad (3.17)$$

where $\hat{p}_1, \hat{p}_1, ..., \hat{p}_M$ should be found by using the method in [71]. If the rows of the channel matrix are ordered into $\mathbf{H}^{(L)}$ and by using $\mathbf{H}^{(L)}$ the matrix $\mathbf{G}_C$ becomes $\mathbf{G}_C^{(L)}$, the average BER of all mobiles becomes

$$P_e = (1/M) \sum_{k=1}^{M} P(p_k / (\sigma_n^2 [\mathbf{G}_C^{(L)}]_{kk}^2)). \quad (3.18)$$

In this case, to minimize the average BER, the optimal order $\bar{L}$ should be found jointly with the power allocation factors $p_1, p_2, ..., p_M$, i.e.,

$$\{\bar{L}, \hat{p}_1, \hat{p}_1, ..., \hat{p}_M\} = \arg\min_{L, p_1, p_2, ..., p_M} [(1/M) \sum_{k=1}^{M} P(p_k / (\sigma_n^2 [\mathbf{G}_C^{(L)}]_{kk}^2))]. \quad (3.19)$$

However, for the above algorithms, the optimal order can only be found by "exhaustive search" method, i.e., all possible orders need to be tested and the one achieving the optimal performance is chosen. Particularly, for the minimum BER ZF nonlinear pre-processing algorithm, the optimal order and the optimal power allocation factors should be found together by "exhaustive search" method. How to design ordering method with reduced complexity for the above algorithms is

still an open question. For the minimum BER ZF nonlinear pre-processing algorithm, the iterative method proposed to find the order and power allocation factors for the ZF DFD in [71] can be applied and this method can reduce the complexity. However, there is no proof of the convergence of this method.

Interestingly, we find by simulation that the "best-first" ordering for the conventional ZF nonlinear pre-processing algorithm introduced in Section 3.4 can achieve very satisfactory performance for the balanced ZF nonlinear pre-processing algorithm and the minimum BER ZF nonlinear pre-processing algorithm. For the balanced ZF nonlinear pre-processing algorithm, which is a special case of the optimal ZF nonlinear pre-processing algorithm under the relative SNR requirement, it has been shown in [83] that the "best-first" ordering achieves the optimal order in all $2 \times N$ systems. For the minimum BER ZF nonlinear pre-processing algorithm, we propose to use first the "best-first" ordering to order the channel matrix, and then determine the power allocation factors using the ordered channel matrix. Simulation in the next section will show that the "best-first" ordering can significantly improve the performance of these algorithms, while its complexity is much lower than that of the "exhaustive search" method.

3.5.1 Simulation Results

3.5.1.1 Performance Improvement Achieved by the "Best-First" Ordering for the Balanced ZF Nonlinear Pre-Processing Algorithm

To show how well the "best-first" ordering can work for the balanced ZF nonlinear pre-processing algorithm, we compare it with the "exhaustive search" ordering. It is noted that the "exhaustive search" ordering is optimal, and it can minimize $\beta \equiv \sum_{k=1}^{M} ([\mathbf{G}_C^{(L)}]_{kk}^2)$. Simulation is used, and 10^6 channel gain matrices are tested. For each realization of the channel both the "best-first" ordering and the "exhaustive search" ordering are used, and the β obtained by the two ordering methods are found. Table 3.1 shows the comparison of the results of the two ordering methods when $M = N = 2,\ldots,6$. The value of the "Percentage of failure

(%)" is the percentage of the cases that the "best-first" ordering fails to find the optimal order, i.e., β obtained by the "best-first" ordering (β_{BF}) is larger than β obtained by the "exhaustive search" ordering (β_{ES}). The average values of β_{BF} and β_{ES}, which are represented by $\overline{\beta}_{BF}$ and $\overline{\beta}_{ES}$ respectively, are also shown. From Table 3.1, it can be seen that the "best-first" ordering achieves the optimal order for most channel realizations when the M and N are small. As it is shown in [83], when $M = N = 2$ the "best-first" ordering is optimal. Although as M and N become larger the percentage of cases such that the "best-first" ordering fails to find the optimal order increases, the average value of the β obtained by the "best-first" ordering is still very close to that obtained by the "exhaustive search" ordering. Therefore, the "best-first" ordering achieves very satisfactory results. It is then justified to expect that the performance when the "best-first" ordering is applied is very close to the performance when the optimal order is used. Next, we will use simulation to verify this anticipation.

Table 3.1 Comparison of the "best-first" ordering and the "exhaustive search" ordering for the balanced ZF nonlinear pre-processing algorithm.

M and N	2	3	4	5	6
Percentage of failure (%)	0	1.57	4.98	10.05	16.36
$\overline{\beta}_{BF}$	8.3207	6.9741	5.6255	5.3905	5.0942
$\overline{\beta}_{ES}$	8.3207	6.9728	5.6225	5.3856	5.0875

The improved performance attained by the "best-first" ordering in the balanced ZF nonlinear pre-processing algorithm is shown in Figure 3.5. A multi-user MIMO system with 4 transmit antennas at the base station and 4 single-antenna mobiles is considered. BPSK modulation is used. We assume the channel matrix $\mathbf{H}$ changes independently from one symbol interval to the next. The SNR used in our simulation is defined by (2.43). It can be seen that using the "best-first" ordering, approximately 5.5 dB gain is obtained for the balanced ZF nonlinear pre-processing algorithm at BER = 10^{-3}. The performance curve of the balanced ZF nonlinear pre-processing algorithm when the "exhaustive search"

ordering method is used is also drawn in Figure 3.5. It can be observed that the "best-first" ordering achieves nearly the same performance as the optimal "exhaustive search" ordering. Hence, it is clear that the "best-first" ordering is very beneficial to the system performance.

3.5.1.2 Performance Improvement Achieved by the "Best-First" Ordering for the Minimum BER ZF Nonlinear Pre-Processing Algorithm

Figure 3.6 and Figure 3.7 show the performance improvement by the "best-first" ordering for the minimum BER ZF nonlinear pre-processing algorithm. A multi-user MIMO system with 4 transmit antennas at the base station and 4 single-antenna mobiles is considered in Figure 3.6, and a multi-user MIMO system with 8 transmit antennas at the base station and 8 single-antenna mobiles is considered in Figure 3.7. BPSK modulation is used. We assume the channel matrix **H** changes independently from one symbol interval to the next. The SNR used in our simulation is defined by (2.43). It can be seen that the "best-first" ordering can significantly improve the performance of the minimum BER ZF nonlinear pre-processing algorithm. For the system with $M = N = 4$, using the "best-first" ordering, approximately 4.1 dB gain is obtained for the minimum BER ZF nonlinear pre-processing algorithm at BER = 10^{-3}. For the system with $M = N = 8$, approximately 4.8 dB gain is obtained using the "best-first" ordering at BER = 10^{-3}.

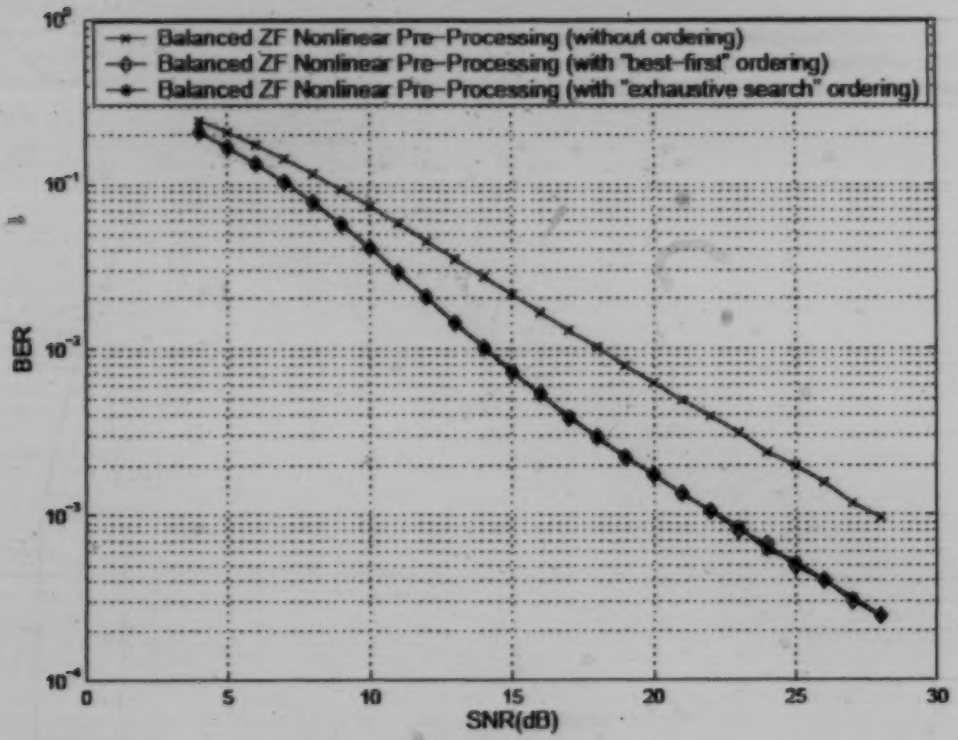


Figure 3.5 Performance comparison of the balanced ZF nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is used (BPSK, $M = N = 4$).

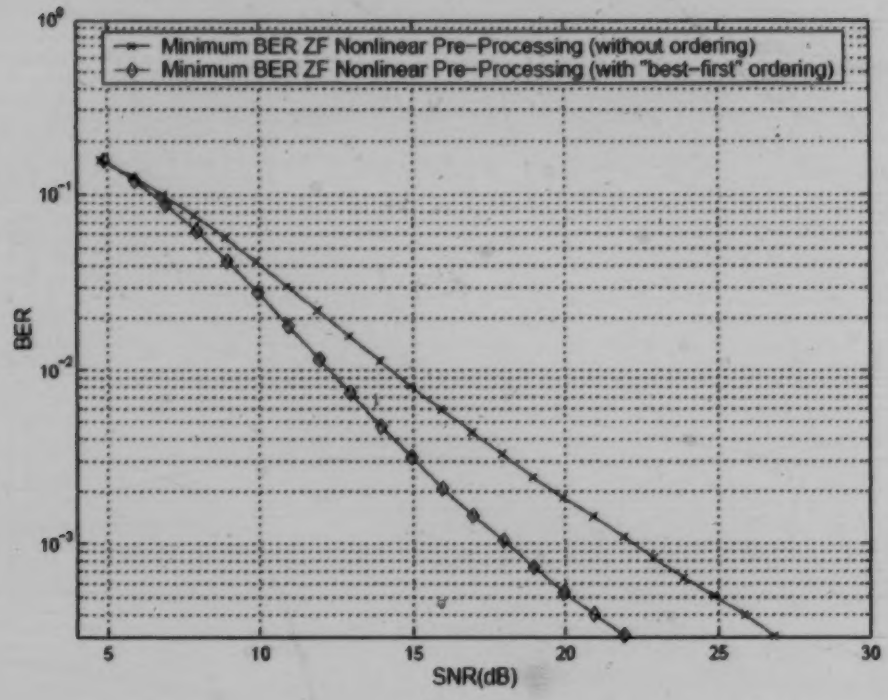


Figure 3.6 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is used (BPSK, $M = N = 4$).

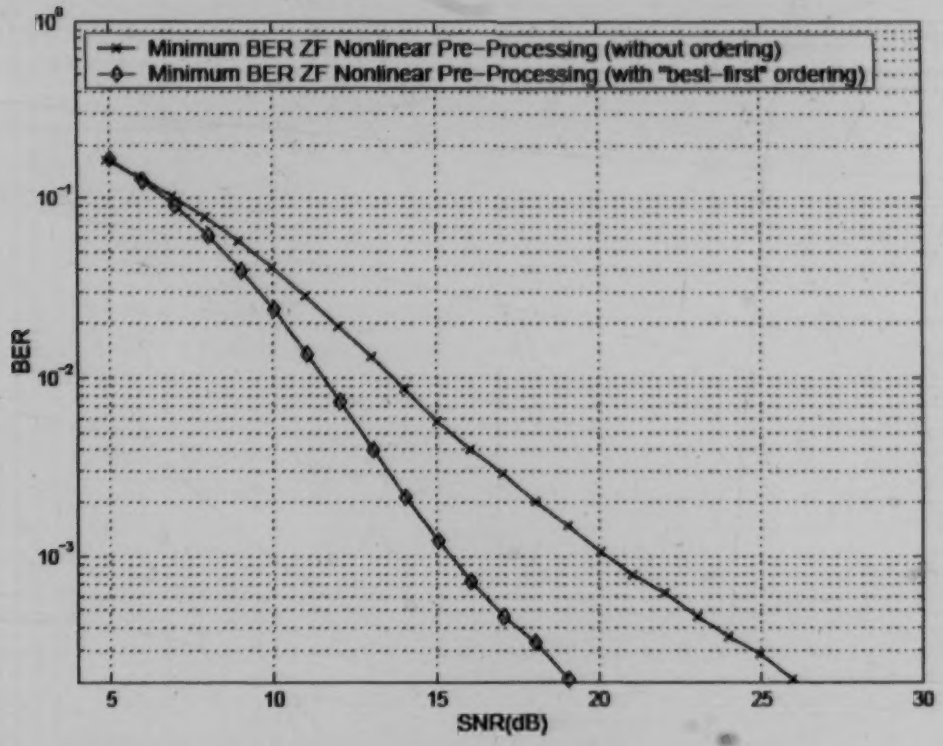


Figure 3.7 Performance comparison of the minimum BER ZF nonlinear pre-processing algorithm when no ordering or the "best-first" ordering is used (BPSK, $M = N = 8$).

Chapter 4 Nonlinear Joint Transmitter-Receiver Processing Algorithms

In this chapter, the downlink transmission of multi-user MIMO systems where the mobiles are equipped with multiple antennas is considered. For this kind of systems, since there is more than one antenna deployed at each mobile, although the mobiles still cannot process signals cooperatively, each mobile can process the multiple signal streams it receives at its own multiple antennas. Therefore, the processing capabilities of the mobiles can be utilized to realize joint Tx-Rx processing algorithms. If designed properly, the joint Tx-Rx processing algorithms can achieve better performance than the Tx pre-processing only algorithms.

Linear joint Tx-Rx processing algorithms, where linear transmitter processing and linear receiver processing are applied at the base station and each mobile, have been proposed in [53]–[56] for the downlink of multi-user MIMO systems with multiple-antenna mobiles, but the processing matrices there have to be found by iterative methods, which leads to high computational complexity. On the other hand, a linear multi-user MIMO decomposition (L-DECOM) technique, which implements joint Tx-Rx processing and provides closed-form solution, is proposed in [57]–[60]. This technique uses linear pre-processing to pre-eliminate the MUI, the interference from the data streams of other users, at the transmitter. It decomposes the multi-user MIMO channel into parallel independent single-user MIMO channels, and then the data streams of each mobile can be seen as passing through a decomposed single-user MIMO channel without any MUI. The

maximum achievable sum rate of multi-user MIMO channel using this L-DECOM technique is analyzed in [57]–[59]. The application of different single-user MIMO processing algorithms to achieve spatial multiplexing in the decomposed single-user MIMO channels is considered in [59] and [60]. With the application of different single-user MIMO spatial layer separation algorithms, joint Tx-Rx processing algorithms for the multi-user MIMO downlink with multiple-antenna mobiles are realized, and closed-form expressions for processing matrices can be derived. However, the performance of these joint Tx-Rx processing algorithms is limited by the linear pre-processing used by the L-DECOM technique.

It has been revealed in Chapter 2 that the nonlinear THP can achieve better performance than linear pre-processing. Therefore, in this chapter we propose a nonlinear multi-user MIMO decomposition (NL-DECOM) technique for the multi-user MIMO downlink. Unlike the L-DECOM technique of [57]–[60], the NL-DECOM technique uses nonlinear pre-processing instead of linear pre-processing to pre-eliminate the MUI at the transmitter. The data streams of each user can also be seen as passing through an equivalent single-user MIMO channel without any MUI, and then the self-interference of each user can be eliminated by single-user MIMO layer separation algorithms. The advantages of the NL-DECOM technique over the L-DECOM technique are analyzed. It will be shown that the NL-DECOM technique can achieve better BER performance and achieve much higher maximum sum rate at high SNR than the L-DECOM technique. Moreover, the restriction on the number of transmit antennas and the restriction on the number of data streams for each mobile in the NL-DECOM technique are less strict than those in the L-DECOM technique, which means the NL-DECOM technique affords higher flexibility in system design. When the NL-DECOM technique is applied, it is found that the mobiles can be ordered properly to further improve the system performance, and the "best-first" ordering method is proposed to improve the overall system performance in the systems, in which all the mobiles have the same number of antennas and use the same self-interference cancellation algorithm.

When the NL-DECOM technique is applied to achieve better performance and reduce complexity of the mobiles, it is desirable to combine the single-user MIMO spatial layer separation algorithms that use the nonlinear THP and linear receiver processing with the NL-DECOM technique. However, since the modulo device in the THP changes the transmitted signal [51], the single-user MIMO spatial layer separation algorithms that employ the THP cannot be combined with the NL-DECOM technique directly. Fortunately, we find that the application of the single-user MIMO spatial layer separation algorithms that use the nonlinear THP and linear receiver processing together with the NL-DECOM technique can be realized by joint design of nonlinear pre-processing for MUI and self-interference suppression. In this regard, we propose the ZF nonlinear joint Tx-Rx processing algorithm and the MMSE nonlinear joint Tx-Rx processing algorithm. The ZF criterion based algorithm, the ZF nonlinear joint Tx-Rx processing algorithm, can ensure the complete elimination of the inter-stream interference, while it does not require the knowledge of the statistics of the transmitted signal and noise. The MMSE criterion based algorithm, the MMSE nonlinear joint Tx-Rx processing algorithm, minimizes the MSE, which includes both the residual inter-stream interference and noise. Hence, it mitigates noise enhancement, and therefore is capable of better performance in comparison to its ZF-based counterpart. In these two algorithms, closed-form expressions for the transmitter and receiver processing matrices are developed to optimize the performance of each mobile. Since the processing capabilities of the mobiles are utilized effectively, the proposed ZF/MMSE nonlinear joint Tx-Rx processing algorithm achieves significantly better performance than the ZF/MMSE nonlinear pre-processing algorithm in Chapter 2. For these two algorithms, it is found that performance can be further improved if the channel matrices of different mobiles are ordered properly. A combined optimal diversity and "best-first" (CODBF) ordering method, which can achieve the optimal order in the maximal diversity order sense, is proposed for the ZF nonlinear joint Tx-Rx processing algorithm. For the MMSE nonlinear joint Tx-Rx processing algorithm the "best-first" ordering method is proposed to perform the ordering.

4.1 System Model and Assumptions

Since the multi-user MIMO systems with multiple-antenna mobiles are the focus of this chapter, we introduce a system model that is different from that in Section 2.1.

We assume the base station has N antennas and there are K mobiles, MS_1, MS_2,…,MS_K, active in this system. MS_k has m_k antennas, and $M = \sum_{k=1}^{K} m_k$ is the total number of antennas at all the mobiles.

Vectors $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_K$ represent the data vectors for MS_1, MS_2, …, MS_K, respectively. We assume l_k ($l_k \le m_k$) independent data streams are transmitted to MS_k, so $\mathbf{a}_k$ is an $l_k \times 1$ vector. The $L \times 1$ ($L = \sum_{k=1}^{K} l_k$) vector $\mathbf{a} \equiv (a_1, a_2, ..., a_L)^T = (\mathbf{a}_1^T, \mathbf{a}_2^T, ..., \mathbf{a}_K^T)^T$ is used to represent the data for all the mobiles. It is assumed that all the data symbols are independent and have unit power, i.e., $\mathbf{R}_{\mathbf{a}} \equiv E[\mathbf{a}\mathbf{a}^H] = \mathbf{I}$.

Since transmitter pre-processing is applied, we assume $\mathbf{x} \equiv (x_1, x_2, ..., x_N)^T$ is the pre-processing output of the data vector $\mathbf{a}$, where $x_1, x_2, ..., x_N$ are the signals transmitted from the N transmit antennas. The received signal vector $\mathbf{y}_k \equiv (y_k^{(1)}, ..., y_k^{(m_k)})^T$ at MS_k is

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \tag{4.1}$$

where $\mathbf{n}_k \equiv (n_k^{(1)}, ..., n_k^{(m_k)})^T$ is the noise vector made up of the noise samples at different receive antennas of MS_k, and $\mathbf{H}_k$ is the $m_k \times N$ channel gain matrix representing the channel between MS_k and the base station. Flat Rayleigh fading is assumed. The entries of $\mathbf{H}_k$ are assumed to be normalized i.i.d. zero-mean complex Gaussian random variables. $n_k^{(1)}, ..., n_k^{(m_k)}$ are assumed to be i.i.d. zero-mean complex Gaussian random variables with the covariance matrix $\mathbf{R}_{\mathbf{n}_k} \equiv E[\mathbf{n}_k \mathbf{n}_k^H] = \sigma_n^2 \mathbf{I}$.

Another way to represent this system is to arrange the received signals at all the receive branches into one larger vector

$$\mathbf{y} \equiv (y_1, y_2, \ldots, y_M)^T = (\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_K^T)^T . \tag{4.2}$$

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{4.3}$$

where

$$\mathbf{n} \equiv (n_1, n_2, \ldots, n_M)^T = (\mathbf{n}_1^T, \mathbf{n}_2^T, \ldots, \mathbf{n}_K^T)^T , \tag{4.4}$$

and

$$\mathbf{H} \equiv [\mathbf{H}_1^T, \mathbf{H}_2^T, \ldots, \mathbf{H}_K^T]^T . \tag{4.5}$$

4.2 Multi-User MIMO Decomposition Techniques

In this section, the L-DECOM technique proposed in [57]–[60] is first briefly described. Then, our novel NL-DECOM technique is proposed, the advantages of the NL-DECOM technique over the L-DECOM technique are discussed, and then the ordering problem for the proposed technique is addressed. Finally, simulation is used to show the advantages of the NL-DECOM technique and the "best-first" ordering method applied in the NL-DECOM technique.

4.2.1 Linear Multi-User MIMO Decomposition Technique

The L-DECOM technique proposed in [57]–[60] uses linear pre-processing to pre-eliminate the MUI at the transmitter. Assuming $\mathbf{T} \equiv [\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_K]$ is the linear pre-processing matrix where $\mathbf{T}_k$ is the linear pre-processing matrix for the data of MS_k, the precoding output is $\mathbf{x} = \mathbf{T}\mathbf{a}$. To pre-eliminate the MUI at the transmitter, the L-DECOM technique requires $\mathbf{H}_j\mathbf{T}_k = \mathbf{0}_{m_j, J_k}$ for $j \neq k$. In other words, the pre-processing matrix $\mathbf{T}$ should make $\mathbf{HT}$ a block diagonal matrix.

$\mathbf{T}_k$ in this technique can be represented as $\mathbf{T}_k = \mathbf{N}_k^{(L)}\mathbf{A}_k^{(L)}$, where $\mathbf{N}_k^{(L)}$ is a matrix whose columns are the orthonormal basis vectors of the null space of $\bar{\mathbf{H}}_k$, the null($\bar{\mathbf{H}}_k$), and $\bar{\mathbf{H}}_k \equiv [\mathbf{H}_1^T, \ldots, \mathbf{H}_{k-1}^T, \mathbf{H}_{k+1}^T, \ldots, \mathbf{H}_K^T]^T$. The $N \times (N - \sum_{i=1, i\neq k}^{K} m_i)$ matrix

$\mathbf{N}_k^{(L)}$ can be found by SVD with respect to $\bar{\mathbf{H}}_k$ [61]. Since with this $\mathbf{N}_k^{(L)}$, $\mathbf{H}_j\mathbf{T}_k$ is a zero matrix for any $j \neq k$, $\mathbf{A}_k^{(L)}$ can be any matrix of size $(N - \sum_{i=1, i\neq k}^{K} m_i) \times l_k$.

Applying the L-DECOM technique, it can be found that the received signal vector at MS_k is $\mathbf{y}_k = \mathbf{H}_k\mathbf{T}\mathbf{a} + \mathbf{n}_k = \mathbf{H}_k\mathbf{N}_k^{(L)}\mathbf{A}_k^{(L)}\mathbf{a}_k + \mathbf{n}_k$. Thus, the data for MS_k can be seen as passing through an equivalent single-user MIMO channel $\mathbf{H}_k\mathbf{N}_k^{(L)}$ without any MUI. Matrix $\mathbf{A}_k^{(L)}$ can be seen as the equivalent transmitter-processing matrix for MS_k, and it can be designed following the same rule for the single-user MIMO systems. Therefore, all kinds of single-user MIMO layer separation algorithms can be adopted to mitigate the self-interference for each user.

4.2.2 Nonlinear Multi-User MIMO Decomposition Technique

The NL-DECOM technique we propose uses nonlinear TH precoder and a linear transmitter-processing matrix to pre-eliminate the MUI at the transmitter. Its structure is shown in Figure 4.1. In this structure, the linear transmitter-processing matrix $\mathbf{T}$ (size $N \times L$) is required to make $\mathbf{HT}$ a block triangular matrix instead of a block diagonal matrix, i.e., $\mathbf{H}_j\mathbf{T}_k = \mathbf{0}_{m_j, l_k}$ for $j > k$. In other words,

$$\mathbf{HT} = \begin{bmatrix} \mathbf{H}_1\mathbf{T}_1 & \mathbf{H}_1\mathbf{T}_2 & \cdots & \mathbf{H}_1\mathbf{T}_K \\ \mathbf{0} & \mathbf{H}_2\mathbf{T}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{H}_{K-1}\mathbf{T}_K \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}_K\mathbf{T}_K \end{bmatrix}. \tag{4.6}$$

Using this pre-processing matrix $\mathbf{T}$, the data vector $\mathbf{a}_k$ is not interfered by $\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}$ but still interfered by $\mathbf{a}_{k+1}, \ldots, \mathbf{a}_K$. The MUI from $\mathbf{a}_j$ ($j > k$) to $\mathbf{a}_k$ is eliminated by the nonlinear THP.

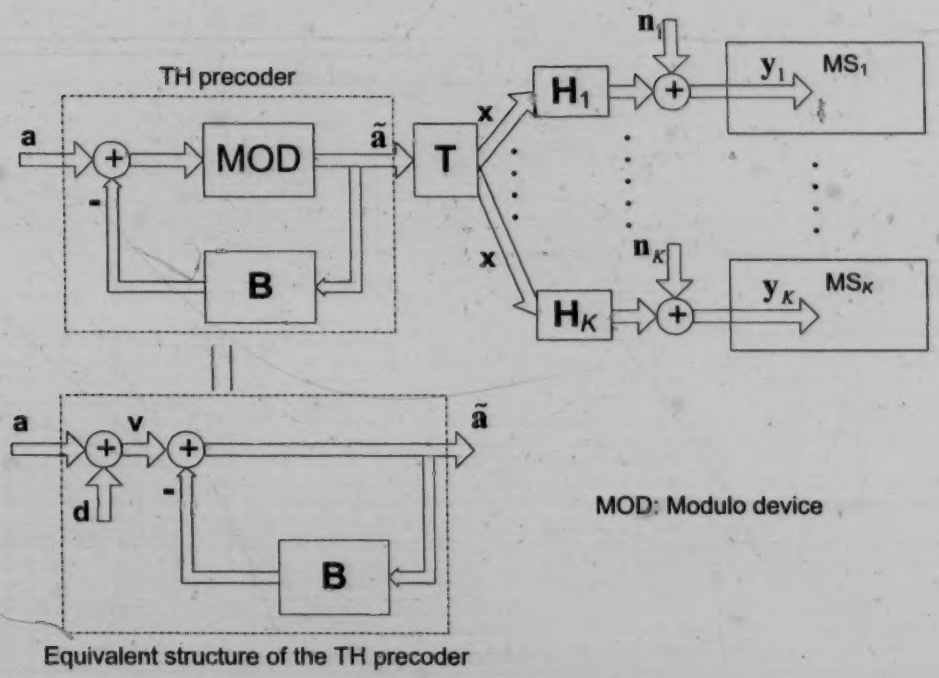


Figure 4.1 Structure of the NL-DECOM technique in the downlink of a multi-user MIMO system.

In Figure 4.1, the block outlined by the dashed line at the transmitter is the TH precoder whose basic idea of operation is described in Chapter 2. Matrix $\mathbf{B}$ (size $L \times L$) works as the feedback filter and it is a strictly upper triangular matrix. The equivalent structure of the TH precoder is also shown in Figure 4.1. The precoding output is

$$\tilde{\mathbf{a}} \equiv (\tilde{\mathbf{a}}_1^T, \tilde{\mathbf{a}}_2^T, ..., \tilde{\mathbf{a}}_K^T)^T = (\mathbf{B}+\mathbf{I})^{-1}\mathbf{v} = (\mathbf{B}+\mathbf{I})^{-1}(\mathbf{a}+\mathbf{d}) \tag{4.7}$$

where $\mathbf{v} \equiv (\mathbf{v}_1^T, \mathbf{v}_2^T, ..., \mathbf{v}_K^T)^T = \mathbf{a}+\mathbf{d}$ is the effective data vector and $\mathbf{d} \equiv (\mathbf{d}_1^T, \mathbf{d}_2^T, ..., \mathbf{d}_K^T)^T$ is the precoding vector.

Now, let us first show how to design the feedback matrix $\mathbf{B}$ to pre-equalize the MUI from $\mathbf{a}_j$ $(j > k)$ to $\mathbf{a}_k$. From (4.3), (4.7), and $\mathbf{x} = \mathbf{T}\tilde{\mathbf{a}}$, one can get

$$\mathbf{y} = \mathbf{HT}\tilde{\mathbf{a}} + \mathbf{n} = \mathbf{HT}(\mathbf{B}+\mathbf{I})^{-1}\mathbf{v} + \mathbf{n}. \tag{4.8}$$

In order to pre-eliminate the MUI at transmitter, matrix $\mathbf{B}$ should satisfy

$$\mathbf{HT}(\mathbf{B}+\mathbf{I})^{-1} = \mathbf{D}_{\mathbf{HT}} \tag{4.9}$$

where $\mathbf{D}_{\mathbf{HT}} \equiv \text{diag}[\mathbf{H}_1\mathbf{T}_1, \mathbf{H}_2\mathbf{T}_2, ..., \mathbf{H}_K\mathbf{T}_K]$. The $\text{diag}[\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_K]$ denotes a block diagonal matrix whose diagonal blocks are $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_K$.

If the number of data streams transmitted to each mobile is equal to the number of antennas at that mobile, i.e., $l_k = m_k$ for k = 1,...,K, the $\mathbf{B}$ satisfying (4.9) can be found to be

$$\begin{aligned}\mathbf{B} &= \mathbf{D}_{\mathbf{HT}}^{-1}\mathbf{HT} - \mathbf{I} \\ &= \begin{bmatrix} \mathbf{0} & (\mathbf{H}_1\mathbf{T}_1)^{-1}\mathbf{H}_1\mathbf{T}_2 & \cdots & \cdots & (\mathbf{H}_1\mathbf{T}_1)^{-1}\mathbf{H}_1\mathbf{T}_K \\ \mathbf{0} & \ddots & (\mathbf{H}_2\mathbf{T}_2)^{-1}\mathbf{H}_2\mathbf{T}_3 & \cdots & (\mathbf{H}_2\mathbf{T}_2)^{-1}\mathbf{H}_2\mathbf{T}_K \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & (\mathbf{H}_{K-1}\mathbf{T}_{K-1})^{-1}\mathbf{H}_{K-1}\mathbf{T}_K \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} \end{bmatrix}\end{aligned} \tag{4.10}$$

Therefore, the received signal vector $\mathbf{y}$ is

$$\mathbf{y} = \mathbf{D}_{\mathbf{HT}}\mathbf{v} + \mathbf{n}, \tag{4.11}$$

and the received signal vector at MS_k is

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{T}_k\mathbf{v}_k + \mathbf{n}_k. \tag{4.12}$$

One can see from (4.11) or (4.12) that all the MUI is pre-eliminated.

If there exists at least one $l_k < m_k$ for k = 1,...,K, both $\mathbf{HT}$ and $\mathbf{D}_{\mathbf{HT}}$ are tall matrices[3], so the matrix $\mathbf{B}$ that satisfies (4.9) cannot be found. In this case, the NL-DECOM technique employs an $l_k \times m_k$ linear receiver-processing matrix $\mathbf{G}_k$ at the mobiles with $l_k < m_k$ to combine the m_k signal streams collected from the antennas into l_k signal streams, i.e., $\mathbf{y}_k' = \mathbf{G}_k\mathbf{y}_k$. Here, $\mathbf{y}_k'$ is used to represent the signal vector at the output of the linear receiver-processing matrix $\mathbf{G}_k$ at MS_k. In this case, we represent the feedback matrix $\mathbf{B}$ as $\mathbf{B} \equiv (\mathbf{B}_1^T, \mathbf{B}_2^T, ..., \mathbf{B}_K^T)^T$, where $\mathbf{B}_k$ is an $l_k \times L$ matrix, and $\mathbf{B}_k$ is found to be

$$\mathbf{B}_k = \begin{cases} [\mathbf{0}_{l_k L_k}, (\mathbf{H}_k\mathbf{T}_k)^{-1}\mathbf{H}_k\mathbf{T}_{k+1}, ..., (\mathbf{H}_k\mathbf{T}_k)^{-1}\mathbf{H}_k\mathbf{T}_K] & \text{if} \quad l_k = m_k \\ [\mathbf{0}_{l_k L_k}, (\mathbf{G}_k\mathbf{H}_k\mathbf{T}_k)^{-1}\mathbf{G}_k\mathbf{H}_k\mathbf{T}_{k+1}, ..., (\mathbf{G}_k\mathbf{H}_k\mathbf{T}_k)^{-1}\mathbf{G}_k\mathbf{H}_k\mathbf{T}_K] & \text{if} \quad l_k < m_k \end{cases} \tag{4.13}$$

---

[3] Tall matrix is defined as a matrix whose number of rows is larger than the number of columns.

where $L_k = \sum_{i=1}^{k} l_i$. Using (4.6)–(4.8) and **B** given by (4.13), one can find that for the mobiles with $l_k = m_k$, the received signal vector is

$$\begin{aligned}\mathbf{y}_k &= \mathbf{H}_k\mathbf{T}\tilde{\mathbf{a}} + \mathbf{n}_k \\ &= \mathbf{H}_k\mathbf{T}_k\tilde{\mathbf{a}}_k + \sum_{i=k+1}^{K}\mathbf{H}_k\mathbf{T}_i\tilde{\mathbf{a}}_i + \mathbf{n}_k \\ &= \mathbf{H}_k\mathbf{T}_k(\mathbf{v}_k - \mathbf{B}_k\tilde{\mathbf{a}}) + \sum_{i=k+1}^{K}\mathbf{H}_k\mathbf{T}_i\tilde{\mathbf{a}}_i + \mathbf{n}_k \\ &= \mathbf{H}_k\mathbf{T}_k\mathbf{v}_k + \mathbf{n}_k\end{aligned} \tag{4.14}$$

which is the same as that in (4.12). Similarly, for the mobiles with $l_k < m_k$, it can be found that

$$\mathbf{y}_k' = \mathbf{G}_k\mathbf{H}_k\mathbf{T}_k\mathbf{v}_k + \mathbf{G}_k\mathbf{n}_k. \tag{4.15}$$

Therefore, in both cases, for each mobile, all the MUI is pre-eliminated at the transmitter.

In the NL-DECOM technique, $\mathbf{T}_k$ that satisfies (4.6) can be represented as $\mathbf{T}_k = \mathbf{N}_k^{(NL)}\mathbf{A}_k^{(NL)}$. $\mathbf{N}_k^{(NL)}$ is a matrix whose columns are the orthonormal basis vectors of null($\tilde{\mathbf{H}}_k$), where $\tilde{\mathbf{H}}_k \equiv [\mathbf{H}_{k+1}^T, ..., \mathbf{H}_K^T]^T$. $\mathbf{N}_k^{(NL)}$ can be found by doing the SVD on $\tilde{\mathbf{H}}_k$, i.e.,

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k\begin{bmatrix}\boldsymbol{\Sigma}_k & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{N}}_k^{(NL)H} \\ \mathbf{N}_k^{(NL)H}\end{bmatrix}. \tag{4.16}$$

The size of $\mathbf{N}_k^{(NL)}$ is $N \times (N - \sum_{i=k+1}^{K} m_i)$. In this manner, (4.12) and (4.15) become

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{N}_k^{(NL)}\mathbf{A}_k^{(NL)}\mathbf{v}_k + \mathbf{n}_k \tag{4.17}$$

and

$$\mathbf{y}_k' = \mathbf{G}_k\mathbf{H}_k\mathbf{N}_k^{(NL)}\mathbf{A}_k^{(NL)}\mathbf{v}_k + \mathbf{G}_k\mathbf{n}_k, \tag{4.18}$$

respectively. Therefore, the data of MS_k can be seen as passing through an equivalent single-user MIMO channel $\mathbf{H}_k\mathbf{N}_k^{(NL)}$, and $\mathbf{A}_k^{(NL)}$ (size $(N - \sum_{i=k+1}^{K} m_i) \times l_k$) can be seen as the transmitter-processing matrix for MS_k. The equivalent single-

user MIMO channels formed using the NL-DECOM technique are shown in Figure 4.2.

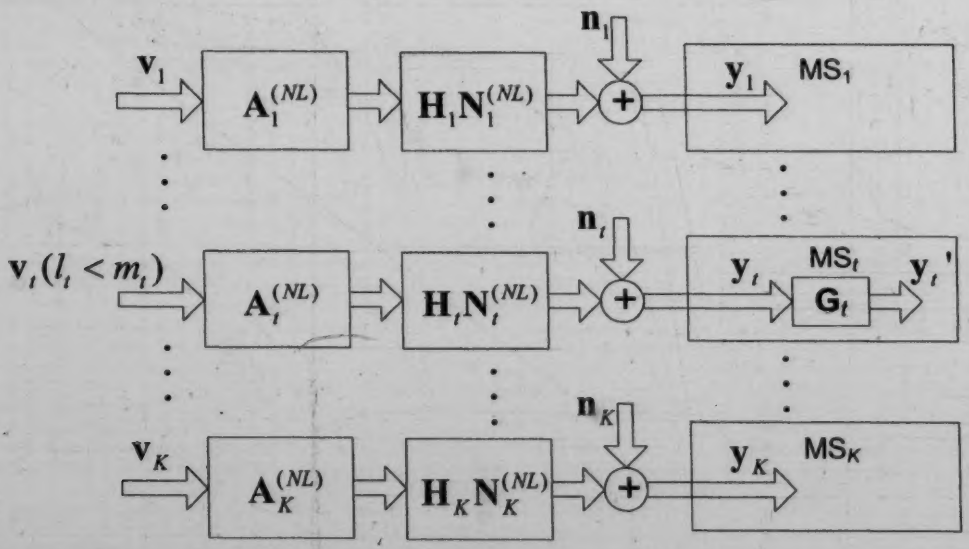


Figure 4.2 The equivalent single-user MIMO channels formed using the NL-DECOM technique.

4.2.3 Advantages of the NL-DECOM Technique over the L-DECOM Technique

4.2.3.1 Better BER Performance

For both the L-DECOM technique of [57]–[60] and the NL-DECOM technique we have proposed in Section 4.2.2, the data for MS_k can be seen as passing through an equivalent single-user MIMO channel. Since the columns of $\mathbf{N}_k^{(L)}$ and $\mathbf{N}_k^{(NL)}$ are orthonormal in these two techniques and it is assumed that the entries of $\mathbf{H}_k$ are normalized i.i.d. zero-mean complex Gaussian random variables, the entries of the equivalent channels $\mathbf{H}_k\mathbf{N}_k^{(L)}$ and $\mathbf{H}_k\mathbf{N}_k^{(NL)}$ are still i.i.d. complex Gaussian random variables with normalized power. For the L-DECOM technique, the size of $\mathbf{H}_k\mathbf{N}_k^{(L)}$ is $m_k \times (N - \sum_{i=1, i\neq k}^{K} m_i)$, while for the NL-DECOM technique,

the size of $\mathbf{H}_k\mathbf{N}_k^{(NL)}$ is $m_k \times (N - \sum_{i=k+1}^{K} m_i)$. Therefore, for MS_k ($k>1$) the NL-DECOM technique realizes an equivalent single-user MIMO channel with more transmit antennas compared to the L-DECOM technique. This can lead to better BER performance, as it will be shown by simulation in Section 4.2.6.1.

4.2.3.2 Higher Maximum Achievable Sum Rate

When the L-DECOM technique is applied, the maximum achievable sum rate of the system is given in [57] as

$$R_L = \max_{\mathbf{\Lambda}} \log_2 |\mathbf{I} + (\mathbf{\Sigma}^{(L)^2}\mathbf{\Lambda})/\sigma_n^2| \tag{4.19}$$

where

$$\mathbf{\Sigma}^{(L)} = \begin{bmatrix} \mathbf{\Sigma}_1^{(L)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2^{(L)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{\Sigma}_K^{(L)} \end{bmatrix} \tag{4.20}$$

and $\mathbf{\Sigma}_k^{(L)}$ is found from the SVD of $\mathbf{H}_k\mathbf{N}_k^{(L)}$, i.e.,

$$\mathbf{H}_k\mathbf{N}_k^{(L)} = [\mathbf{U}_k^{(L)} \quad \tilde{\mathbf{U}}_k^{(L)}] \begin{bmatrix} \mathbf{\Sigma}_k^{(L)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_k^{(L)H} \\ \tilde{\mathbf{V}}_k^{(L)H} \end{bmatrix}. \tag{4.21}$$

Assume the size of $\mathbf{\Sigma}^{(L)}$ is $D_L \times D_L$, the matrix $\mathbf{\Lambda}$ in (4.19) is a $D_L \times D_L$ diagonal matrix whose diagonal elements λ_i decide the power allocated to each date stream, i.e., $\sum_{i=1}^{D_L} \lambda_i = P$. The optimal $\mathbf{\Lambda}$ that achieves the maximum achievable sum rate can be found by the well-known water-filling process [85]. At high SNR, the transmitted power allocated to different data streams is almost the same.

When the NL-DECOM technique is applied, the maximum achievable sum rate in the system will be found in the following. First, the SVD of $\mathbf{H}_k\mathbf{N}_k^{(NL)}$ is found to be

$$\mathbf{H}_k\mathbf{N}_k^{(NL)} = [\mathbf{U}_k^{(NL)} \quad \tilde{\mathbf{U}}_k^{(NL)}] \begin{bmatrix} \mathbf{\Sigma}_k^{(NL)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_k^{(NL)H} \\ \tilde{\mathbf{V}}_k^{(NL)H} \end{bmatrix}. \tag{4.22}$$

Assuming $\mathbf{V}_k^{(NL)}$ is used as the linear pre-processing matrix $\mathbf{A}_k^{(NL)}$ and $\mathbf{U}_k^{(NL)H}$ is used as the linear receiver-processing matrix for MS_k, the received signal vector that will go through the modulo and decision devices is

$$\mathbf{y}_k' = \boldsymbol{\Sigma}_k^{(NL)}\mathbf{v}_k + \mathbf{U}_k^{(NL)H}\mathbf{n}_k. \tag{4.23}$$

Using the same analysis as in [76] and [84], one can find that the maximum achievable sum rate of the system is

$$R_{NL} = \max_{\{p_i\}} \sum_{i=1}^{D_{NL}} \Re(\sqrt{3p_i/2}, \sigma_n/\sigma_i^{(NL)}) \tag{4.24}$$

where $\sigma_i^{(NL)}$ represents the ith diagonal element of $\boldsymbol{\Sigma}^{(NL)}$, and

$$\boldsymbol{\Sigma}^{(NL)} = \begin{bmatrix} \boldsymbol{\Sigma}_1^{(NL)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2^{(NL)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\Sigma}_K^{(NL)} \end{bmatrix} \tag{4.25}$$

is a $D_{NL} \times D_{NL}$ matrix. p_i decides the power allocated to the ith date stream, i.e., $\sum_{i=1}^{D_{NL}} p_i = P$.

$$\Re(\sqrt{3p_i/2}, \sigma_n/\sigma_i^{(NL)}) = 2\log_2(2\sqrt{3p_i/2}) - h(\mathrm{MOD}_{2\sqrt{3p_i/2}}(n/\sigma_i^{(NL)})) \tag{4.26}$$

where $h(\cdot)$ denotes the differential entropy [85], $\mathrm{MOD}_A(\cdot)$ denotes the modulo operation that constrains its output to the interval $[-A/2, A/2)\times[-jA/2, jA/2)$, and n is a zero-mean complex Gaussian random variable with variance σ_n^2. Since finding the maximum achievable sum rate from (4.24) is in general a non-convex problem, the power allocation factors p_i that achieve the maximum achievable sum rate are not easy to find in closed-form. In this thesis, we only consider the maximum achievable sum rate when the SNR is high. In this case, the transmitted power is allocated almost equally to all data streams [76]. With the high SNR assumption, using the same approach as that in [76] and [84] it can be found that the maximum achievable sum rate is

$$R_{NL} = \sum_{i=1}^{D_{NL}} \log_2(\tilde{p}_i \sigma_i^{(NL)2}/\sigma_n^2) - D_{NL}\log_2(\pi e/6) \tag{4.27}$$

where $\tilde{p}_i = P / D_{NL}$.

The maximum achievable sum rate of the multi-user MIMO channel at high SNR when the NL-DECOM technique and the L-DECOM technique are used will be compared by simulation in Section 4.2.6 (see Figure 4.5 and Figure 4.6). It will be found that the NL-DECOM technique outperforms the L-DECOM technique significantly.

4.2.3.3 Higher Flexibility in System Design

When the L-DECOM technique is applied, it has been found that the number of transmit antennas should satisfy [59]

$$N > \max\{\sum_{i=1,i\neq k}^{K} m_i, k = 1,2,...,K\}. \tag{4.28}$$

Also, the size of the equivalent single-user MIMO channel decides that the number of data streams for each mobile, the l_k, should satisfy

$$l_k \leq \min\{m_k, N - \sum_{i=1,i\neq k}^{K} m_i\}. \tag{4.29}$$

When the NL-DECOM technique is applied, it can be found that the number of transmit antennas should satisfy

$$N > \sum_{i=2}^{K} m_i. \tag{4.30}$$

Clearly, this constraint is looser than or the same as the constraint in (4.28). From (4.30) one can see that when the numbers of antennas at different mobiles are unequal, choosing the mobile with the largest number of antennas as MS_1 can loosen the constraint on N. The size of the equivalent single-user MIMO channel and the constraint on N in (4.30) decide that the number of data streams for each mobile should satisfy

$$\begin{aligned} &l_k \leq m_k, \quad k = 2,...,K, \\ &l_1 \leq \min\{m_1, N - \sum_{i=2}^{K} m_i\}. \end{aligned} \tag{4.31}$$

Therefore, the constraint on l_1 is the same as that for the L-DECOM technique, while the constraint on $l_k (k = 2,..., K)$ is looser than or the same as that for the L-DECOM technique.

Based on the above analysis, we can draw the conclusion that the restriction on the number of transmit antennas and the restriction on the number of data streams for each mobile in the NL-DECOM technique are less strict than those in the L-DECOM technique, which means the NL-DECOM technique affords higher flexibility in system design. Two examples are given below to show the higher flexibility of the NL-DECOM technique.

Example 4.1. Let us consider a multi-user MIMO system, in which $N = 5$ transmit antennas are deployed at the base station and there are 2 mobiles with 2 antennas each and 1 mobile with 3 antennas that need to communicate with the base station. Since $\max\{\sum_{i=1, i\neq k}^{3} m_i, k = 1,2,3\} = 5$, from (4.28) one can see that if the L-DECOM technique is applied, the 3 mobiles cannot be accommodated simultaneously. Only 2 out of the 3 mobiles can be served at the same time. For the same system, if we choose $m_1 = 3$ and $m_2 = m_3 = 2$, since $\sum_{i=2}^{3} m_i = 4 < N$, from (4.30) one can see that the 3 mobiles can be accommodated simultaneously using the NL-DECOM technique.

Example 4.2. Let us consider a multi-user MIMO system, in which $N = 5$ and there are 3 mobiles with 2 antennas each. In this case, the 3 mobiles can be accommodated simultaneously by using both the L-DECOM technique and the NL-DECOM technique. From (4.29), one can find that when the L-DECOM technique is applied, only one stream of data can be transmitted to each mobile, i.e., $l_k = 1$ $(k = 1,2,3)$. However, when the NL-DECOM technique is applied, from (4.31) one can find that $l_1 = 1$, $l_2 \leq 2$, and $l_3 \leq 2$ are required, which means that two streams of data can be transmitted to MS_2 and MS_3.

4.2.4 Ordering in the NL-DECOM Technique

From Section 4.2.1, one can see that when the L-DECOM technique is applied, the size of the equivalent single-user MIMO channel for each mobile does not depend on the mobile index k. In other words, choosing which mobile to be MS_1, MS_2, ..., and MS_K does not influence the system performance. However, when the NL-DECOM technique is applied, it is clear that the mobile's index influences the size of the equivalent single-user MIMO channel for that mobile. Therefore, if the mobiles are ordered properly, i.e., MS_1, MS_2, ..., and MS_K are chosen suitably, better performance could be achieved.

An ordering solution for the mobiles in the NL-DECOM technique should take account of the number of antennas at each mobile and the self-interference cancellation algorithms applied for each mobile. Moreover, there exist different ordering schemes designed to achieve different targets. Therefore, it is hard to find a general rule to design the ordering scheme. We assume here that overall system performance improvement is our target. We also assume that all the mobiles have the same number of antennas, and the same self-interference cancellation algorithm is used for each mobile. In this case, the "best-first" ordering method described in Section 3.3 can be used to improve the performance of NL-DECOM technique. Its basic operation is described below.

The "best-first" ordering method orders the mobiles by choosing MS_1, MS_2, to MS_K iteratively. To find MS_i, all the unordered mobiles (the mobiles who have not been chosen as MS_1,...,MS_{i-1}) are tested. The mobile that has the best performance in this position will be chosen as MS_i. (The performance can be measured by the SNR or SINR depending the self-interference cancellation algorithm applied.)

Under our assumptions, one can find that if $k < j$, the performance of MS_k is worse than the performance of MS_j, since the equivalent single-user MIMO channel of MS_k has smaller size than that of MS_j. Therefore, the mobile with smaller index affects the overall system performance more significantly. On the other hand, the "best-first" ordering method always tries to choose the mobile with the best performance to be the mobile with the smallest index. Hence, it can

improve the overall system performance. Simulation results in Section 4.2.6 will show that the "best-first" ordering method can significantly improve the overall system performance when the NL-DECOM technique is applied.

4.2.5 Self-Interference Cancellation Algorithms for the NL-DECOM Technique

It has been shown that an equivalent single-user MIMO channel is realized for each mobile when the NL-DECOM technique is applied (see (4.17) and (4.18)). The SVD plus power allocation approach proposed in Section 4.2.3.2 realizes the maximum achievable sum rate of this system. To apply the NL-DECOM technique in practical systems, a spatial layer separation algorithm is needed for each mobile to cancel the self-interference in its decomposed equivalent single-user MIMO channel. Performing SVD on each equivalent single-user MIMO channel and then applying the bit-loading algorithms originally developed for the digital subscriber line (DSL) systems such as [86] can be seen as a method to approach the maximum achievable sum rate. However, it has been shown in [87] that water-filling, the optimal power allocation scheme for single-user MIMO channels, does not give appreciable gain in capacity at high SNR region. Moreover, the application of bit-loading algorithms designed for DSL systems to MIMO systems may be questioned since the uncoded BER regions of these two systems that are of interest are different [88]. Additionally, the bit-loading algorithms will increase the computational complexity of the transmitter. Considering the above factors, we are interested in the spatial layer separation algorithms without bit-loading that can be combined with the NL-DECOM technique. With this assumption, we find that all the spatial layer separation algorithms that do not need statistical properties of the transmitted data vector, such as the ZF criterion based linear/nonlinear detection, pre-processing and joint Tx-Rx processing algorithms [18], [30], [40], [43], [89], [90], and the SVD-based algorithms used in [59], [91], [92], can be used directly for each mobile when the NL-DECOM technique is applied. When using the MMSE-criterion-based layer separation algorithms, one must know the covariance matrix of the effective data

vector $\mathbf{v}_k$, namely, the $\mathbf{R}_{\mathbf{v}_k} \equiv \mathrm{E}[\mathbf{v}_k \mathbf{v}_k^H]$, which is difficult to determine. In this thesis, for the purpose of comparison, we assume $\mathbf{R}_{\mathbf{v}_k} = \mathbf{I}$ and use the MMSE linear joint Tx-Rx processing algorithm [91], [92] as the self-interference cancellation algorithm when the NL-DECOM technique is applied. The simulation result in Section 4.2.6 (Figure 4.4) will show that even with this impractical assumption the NL-DECOM technique still outperforms the L-DECOM technique significantly. The MLD [30]–[32] can also be utilized as self-interference cancellation algorithm when the NL-DECOM technique is applied. However, one should note that the MLD should operate on the constellation of $\mathbf{v}_k$, which can be seen as an extended version of the modulation constellation of $\mathbf{a}_k$ [51], so the computational complexity is increased. The NL-DECOM technique uses nonlinear pre-processing to pre-eliminate the MUI, so it is normal to consider whether we can apply the single-user MIMO layer separation algorithms that employ the THP and linear receiver processing, since such a configuration can effectively utilize the processing capabilities of the transmitter and receiver and does not need to increase the complexity much. However, due to the modulo operation applied, one can find that this type of single-user MIMO layer separation algorithms cannot be applied directly. The application of the single-user MIMO layer separation algorithms employing the THP and linear receiver processing together with the NL-DECOM technique requires joint design of nonlinear pre-processing for MUI and self-interference suppression. We will solve this design problem in Section 4.3.

4.2.6 Simulation Results

4.2.6.1 BER Performance Comparison of the NL-DECOM and L-DECOM Techniques

First, the BER performance of the NL-DECOM technique and the L-DECOM technique is compared. 16QAM with Gray encoding [74] is used, and the BER is the average BER of all the mobiles. It is assumed that the base station has 6 antennas, and there are 3 mobiles each deployed with 2 antennas communicating with the base station. The flat Rayleigh fading channel is assumed to change

independently from one symbol interval to the next. The SNR used in our simulation is defined by (2.43).

In Figure 4.3, the ZF linear pre-processing (LP) algorithm [40], [43] and the precoded ZF DFD algorithm [89], [90] are used as the self-interference cancellation algorithms when the proposed NL-DECOM technique or the L-DECOM technique is applied. Also, for the NL-DECOM technique, the "best-first" ordering method proposed in Section 4.2.4 is applied to improve the overall performance. It is clear that the proposed NL-DECOM technique can achieve significantly better performance than the L-DECOM technique and the "best-first" ordering method is very beneficial for the system performance. When the ZF LP algorithm is used, at BER = 10^{-3}, the NL-DECOM technique achieves approximately 3 dB gain over the L-DECOM technique, and the NL-DECOM technique with "best-first" ordering achieves approximately 6.5 dB gain over the L-DECOM technique. When the precoded ZF DFD algorithm is used, at BER = 10^{-3}, the NL-DECOM technique achieves approximately 1.5 dB gain over the L-DECOM technique, while the NL-DECOM technique with "best-first" ordering achieves approximately 3.5 dB gain over the L-DECOM technique.

Figure 4.4 shows performance comparison of the proposed NL-DECOM technique and the L-DECOM technique when the MMSE linear joint Tx-Rx processing algorithm [91], [92] is used as the self-interference cancellation algorithm. It is found that the NL-DECOM technique outperforms the L-DECOM technique significantly. At BER = 10^{-3}, the NL-DECOM technique achieves approximately 3.5 dB gain over the L-DECOM technique, while the NL-DECOM technique with "best-first" ordering achieves approximately 6.5 dB gain over the L-DECOM technique.

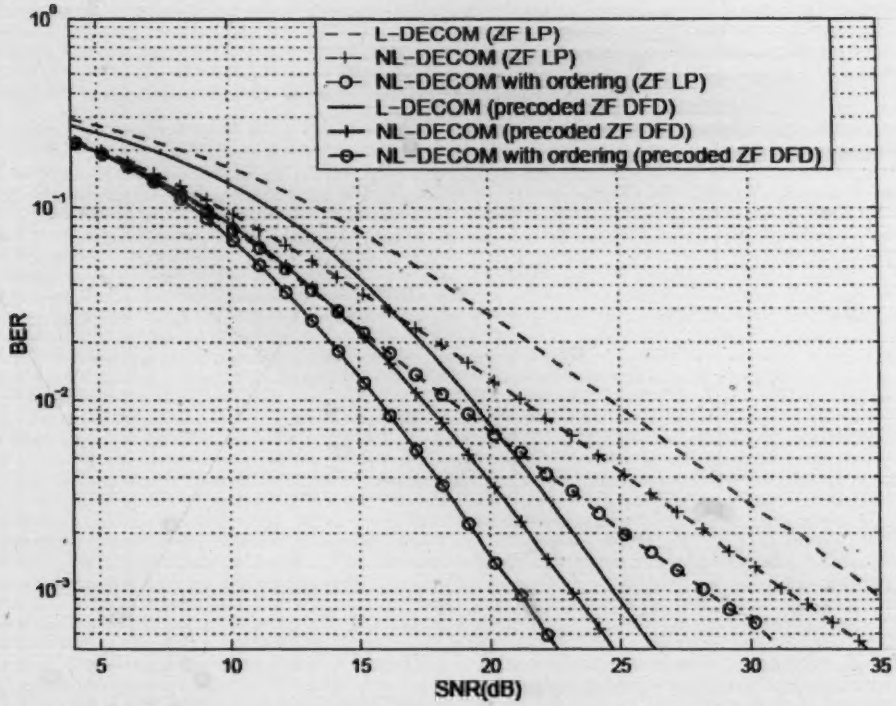


Figure 4.3 Performance comparison of the NL-DECOM and the L-DECOM techniques when the ZF LP or the precoded ZF DFD is used to cancel self-interference (16QAM, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$).

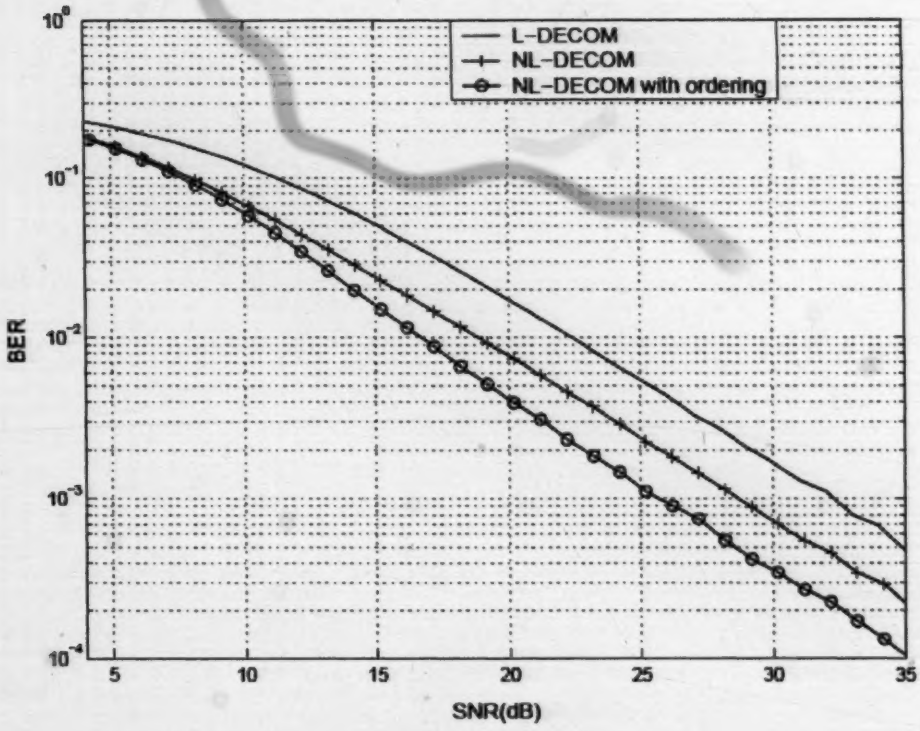


Figure 4.4 Performance comparison of the NL-DECOM and the L-DECOM techniques when the MMSE linear joint Tx-Rx processing algorithm is used to cancel self-interference (16QAM, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$).

4.2.6.2 Comparison of the Maximum Achievable Sum Rate when the NL-DECOM and L-DECOM Techniques are Used

Now, the maximum achievable sum rates of a multi-user MIMO system when the NL-DECOM and the L-DECOM techniques are used are compared. The maximum achievable sum rate of a multi-user MIMO system when the L-DECOM technique is applied is found by the SVD plus water-filling approach as described in Section 4.2.3.2. The maximum achievable sum rate achieved by the NL-DECOM technique results from the SVD plus power allocation approach as shown in Section 4.2.3.2.

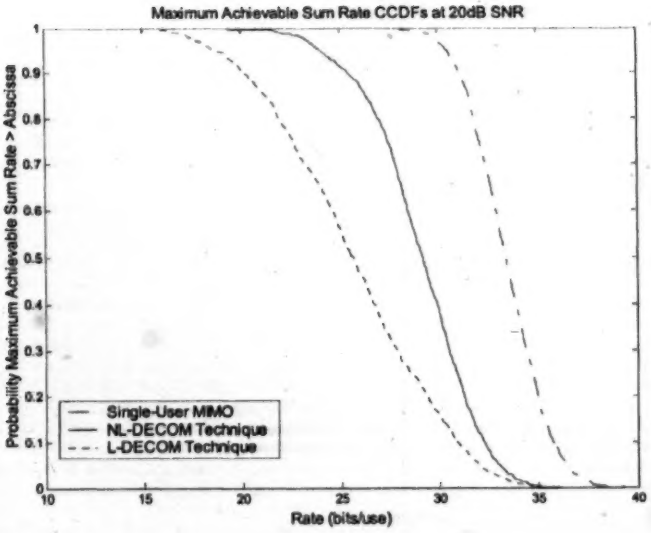
In our simulation, the flat Rayleigh fading channel is assumed to change independently from one symbol interval to the next. The SNR used in this section is defined as

$$\text{SNR} \equiv \bar{E}_{tr} / N_0 \tag{4.32}$$

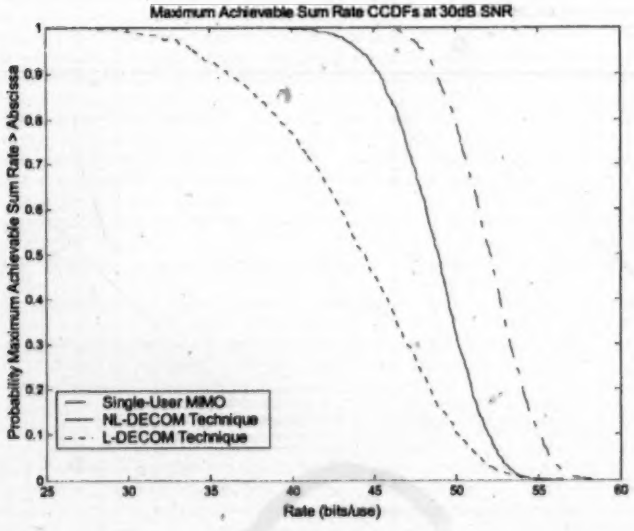
where $\bar{E}_{tr} = \mathrm{E}[\| \mathbf{x} \|^2]$ is the average total transmitted energy per symbol interval, and N_0 is the one-sided noise power spectral density.

Figure 4.5 shows the comparison of the complementary cumulative distribution functions (CCDFs) of the maximum achievable sum rate of a multi-user MIMO system when the NL-DECOM technique and the L-DECOM technique are used. It is assumed that there are 6 transmit antennas at the base station and 3 mobiles each with 2 antennas ($N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$) in the multi-user MIMO system. The CCDF of the capacity of a single-user 6×6 MIMO channel is also shown. One thousand random channel realizations are used to get the result. The SNR is assumed to be 20 dB in Figure 4.5 (a) and 30 dB in Figure 4.5 (b), respectively. The maximum achievable sum rate when the L-DECOM technique is used is generated using (4.19), while the maximum achievable sum rate when the NL-DECOM technique is used is generated based on the high SNR assumption as in (4.24). One can see that the NL-DECOM technique outperforms the L-DECOM technique significantly.

Figure 4.6 compares the average maximum achievable sum rate for one thousand random channel realizations in multi-user MIMO channels with different sizes when the NL-DECOM technique and the L-DECOM technique are used. We assume there are two antennas at each mobile, and the number of transmit antennas at the base station is equal to the total number of receive antennas ($N = M = 4,6,8,10$). The average capacity of the single-user $M \times N$ MIMO channel is also shown. The SNR is assumed to be 20 dB in Figure 4.6 (a) and 30 dB in Figure 4.6 (b), respectively. The advantage of the NL-DECOM technique over the L-DECOM technique can be clearly seen.

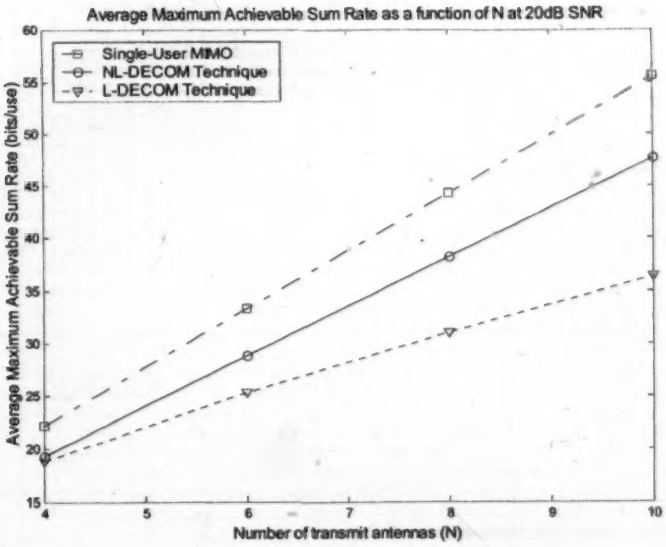


(a)

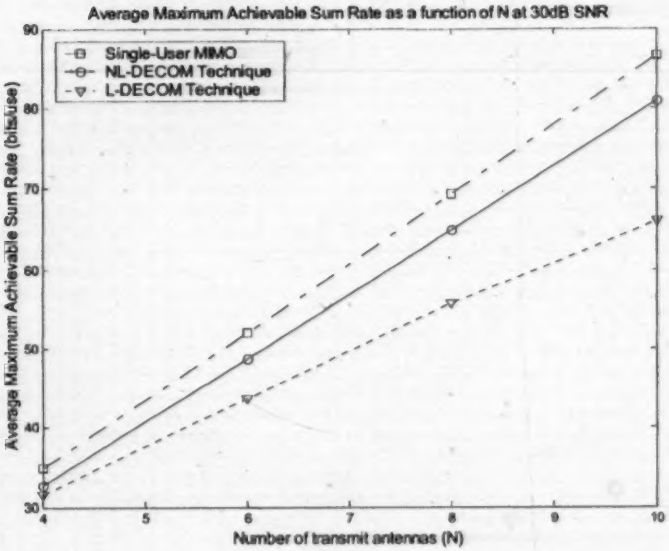


(b)

Figure 4.5 Complementary cumulative distribution functions of the maximum achievable sum rate of a multi-user MIMO channel with $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$. SNR = (a) 20 dB and (b) 30 dB.


Average Maximum Achievable Sum Rate as a function of N at 20dB SNR
Single-User MIMO
NL-DECOM Technique
L-DECOM Technique
Average Maximum Achievable Sum Rate (bits/use)
Number of transmit antennas (N)

(a)


Average Maximum Achievable Sum Rate as a function of N at 30dB SNR
Single-User MIMO
NL-DECOM Technique
L-DECOM Technique
Average Maximum Achievable Sum Rate (bits/use)
Number of transmit antennas (N)

(b)

Figure 4.6 Average maximum achievable sum rate of different multi-user MIMO channels ($M = N = 4, 6, 8, 10$, and $m_k = 2$). SNR = (a) 20 dB and (b) 30 dB.

4.3 ZF/MMSE Nonlinear Joint Tx-Rx Processing Algorithm

In this section, we will study how to design nonlinear joint Tx-Rx processing algorithms that employ nonlinear THP at the transmitter and linear receiver processing at the mobile for the downlink of multi-user MIMO systems with multiple-antenna mobiles. Differently from the NL-DECOM technique in Section 4.2, in the nonlinear joint Tx-Rx processing algorithms proposed in this section the THP at the transmitter suppresses both the MUI and the self-interference. The introduction of these algorithms overcomes the problem that the single-user MIMO spatial layer separation algorithms employing the THP and linear receiver processing cannot be combined with the NL-DECOM technique as we have pointed out in Section 4.2.5.

The ZF and MMSE criteria are used to design the nonlinear joint Tx-Rx processing algorithms, and the resulting algorithms are called the ZF nonlinear joint Tx-Rx processing algorithm and the MMSE nonlinear joint Tx-Rx processing algorithm, respectively. The algorithms in this section effectively utilize the processing capabilities of the base station and the mobiles, while the complexity of the mobiles is kept low since linear receiver processing is applied. They can achieve better performance than with the ZF/MMSE nonlinear pre-processing algorithm or the NL-DECOM technique based joint Tx-Rx processing algorithms with similar or even higher computational complexity.

In this section, the structure of the ZF/MMSE nonlinear joint Tx-Rx processing algorithm is first described in Section 4.3.1. Then the ZF nonlinear joint Tx-Rx processing algorithm is proposed in Section 4.3.2, and the MMSE nonlinear joint Tx-Rx processing algorithm is designed in Section 4.3.3.

4.3.1 Structure of the ZF/MMSE Nonlinear Joint Tx-Rx Processing Algorithm

The system model and assumptions described in Section 4.1 will still be used in this section. For simplicity, here we assume the number of data streams transmitted to each mobile is equal to the number of antennas of that mobile, i.e.,

$l_k = m_k$. Therefore, the total number of data streams transmitted is $M = \sum_{k=1}^{K} m_k$, and $N \geq M$ is assumed for this system.

The structure of the ZF/MMSE nonlinear joint Tx-Rx processing algorithm for the downlink of a multi-user MIMO system with multiple-antenna mobiles is shown in Figure 4.7. In this algorithm, nonlinear pre-processing is applied at the transmitter, while linear receiver processing is applied at each mobile. The principle of the TH precoder employed at the transmitter has been introduced in Section 2.2. The feedback matrix $\mathbf{B}$ is assumed to be a unit upper triangular matrix and is represented as

$$\mathbf{B} \equiv (\mathbf{B}_1^T, \mathbf{B}_2^T, ..., \mathbf{B}_K^T)^T, \tag{4.33}$$

where $\mathbf{B}_k$ is an $m_k \times M$ matrix, or

$$\mathbf{B} \equiv \begin{bmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,K} \\ \mathbf{0} & \mathbf{B}_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{B}_{K-1,K} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{K,K} \end{bmatrix} \tag{4.34}$$

where $\mathbf{B}_{j,k}$ is an $m_j \times m_k$ matrix and $\mathbf{B}_{k,k}$ is still a unit upper triangular matrix.

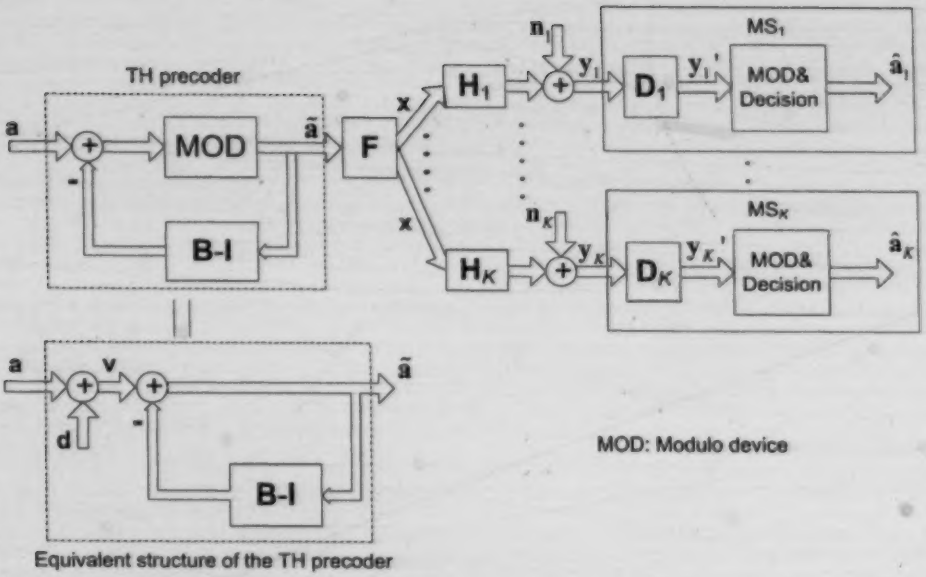


Figure 4.7 Structure of the ZF/MMSE nonlinear joint Tx-Rx processing algorithm.

It is noted that transmitter pre-processing structure of Figure 4.7 is the same as that of the nonlinear pre-processing algorithms described in Section 2.3. The difference is at the receiver that at mobile k, the signal streams collected from the multiple antennas are first processed by a linear receiver-processing matrix $\mathbf{D}_k$ before going into the modulo and decision device. If the receiver-processing matrices $\mathbf{D}_1,...,\mathbf{D}_K$ are all diagonal, this structure is equivalent to the structure of the nonlinear pre-processing algorithms described in Chapter 2. However, in the ZF/MMSE nonlinear joint Tx-Rx processing algorithm, $\mathbf{D}_1,...,\mathbf{D}_K$ are no longer diagonal matrices. This change takes into consideration the processing capabilities of the mobiles, so the ZF/MMSE nonlinear joint Tx-Rx processing algorithm can lead to significantly better performance. However, with this change the design of the processing matrices at the transmitter and the receiver becomes a different and much more challenging problem.

In order to facilitate the representation, we introduce a block diagonal matrix

$$\mathbf{D} \equiv \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{D}_K \end{bmatrix}, \tag{4.35}$$

whose diagonal blocks are the processing matrices at all the mobiles. One can see that the feedforward matrix $\mathbf{F}$ can be represented as $\mathbf{F} \equiv [\mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_K]$ where $\mathbf{F}_k$ is an $N \times m_k$ matrix and can be seen as the linear pre-processing matrix for $\tilde{\mathbf{a}}_k$, the precoding output of MS_k's data streams.

For this system, the signal vector that will go into the modulo and decision devices at the receiver end can be represented as

$$\mathbf{y}' = \mathbf{Dy} = \mathbf{DHFB}^{-1}\mathbf{v} + \mathbf{Dn} = \mathbf{DHFB}^{-1}\mathbf{v} + \bar{\mathbf{n}}. \tag{4.36}$$

where $\bar{\mathbf{n}} \equiv (\bar{\mathbf{n}}_1^T, \bar{\mathbf{n}}_2^T, ..., \bar{\mathbf{n}}_K^T)^T = \mathbf{Dn}$, and $\bar{\mathbf{n}}_k = \mathbf{D}_k\mathbf{n}_k$.

Since the linear pre-processing matrix $\mathbf{F}$ influences the transmitted power, transmitted power constraint has to be applied. We assume $\mathrm{tr}\{\mathbf{FF}^H\} = M$, so from the discussion in Section 2.3, we know that the matrix $\mathbf{F}$ will not change the power of the precoder output vector. We also assume that

$$\mathrm{tr}\{\mathbf{F}_k\mathbf{F}_k^H\} = m_k, \tag{4.37}$$

so the power allocated to each mobile is proportional to the number of data streams transmitted. The design of the ZF/MMSE nonlinear joint Tx-Rx processing algorithm according to other power allocation can be realized using the same approach discussed in Section 4.3.2 and Section 4.3.3.

4.3.2 ZF Nonlinear Joint Tx-Rx Processing Algorithm

In this section, the ZF criterion is used to design processing matrices $\mathbf{F}$, $\mathbf{D}$, and $\mathbf{B}$ for the structure shown in Figure 4.7. The ZF criterion requires all the inter-stream interference should be completely eliminated. Therefore,

$$\mathbf{DHFB}^{-1} = \mathbf{I} \tag{4.38}$$

is required.

To find the solution for processing matrices **F**, **D**, and **B**, first, let us study what kind of matrix **F** can satisfy (4.38). One can see that the ZF requirement of (4.38) can also be written as

$$\mathbf{HF} = \mathbf{D}^{-1}\mathbf{B}. \tag{4.39}$$

Since **B** is assumed to be a unit upper triangular matrix and **D** is a block diagonal matrix, clearly, to satisfy (4.39), the linear pre-processing matrix **F** should be able to make **HF** a block upper triangular matrix, and hence

$$\mathbf{H}_j\mathbf{F}_k = \mathbf{0}_{m_j, m_k}, \qquad j > k \tag{4.40}$$

is required. Therefore,

$$\mathbf{HF} = \begin{bmatrix} \mathbf{H}_1\mathbf{F}_1 & \mathbf{H}_1\mathbf{F}_2 & \cdots & \mathbf{H}_1\mathbf{F}_K \\ \mathbf{0} & \mathbf{H}_2\mathbf{F}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{H}_{K-1}\mathbf{F}_K \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}_K\mathbf{F}_K \end{bmatrix}. \tag{4.41}$$

Equation (4.40) means the columns of $\mathbf{F}_k$ should be in the null space of $\bar{\mathbf{H}}_k$, where

$$\bar{\mathbf{H}}_k \equiv [\mathbf{H}_{k+1}^T, \ldots, \mathbf{H}_K^T]^T. \tag{4.42}$$

Therefore, $\mathbf{F}_k$ that satisfies (4.40) can be represented as $\mathbf{F}_k = \mathbf{N}_k\mathbf{A}_k$. $\mathbf{N}_k$ is an $N \times (N - \sum_{i=k+1}^{K} m_i)$ matrix whose columns are the orthonormal basis vectors of null($\bar{\mathbf{H}}_k$). $\mathbf{A}_k$ can be any matrix of size $(N - \sum_{i=k+1}^{K} m_i) \times m_k$. $\mathbf{N}_k$ can be found by doing the SVD [61] with respect to $\bar{\mathbf{H}}_k$, i.e.,

$$\bar{\mathbf{H}}_k = \tilde{\mathbf{U}}_k \begin{bmatrix} \boldsymbol{\Sigma}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{N}}_k^H \\ \mathbf{N}_k^H \end{bmatrix}. \tag{4.43}$$

Therefore, for $\mathbf{F}_k = \mathbf{N}_k\mathbf{A}_k$, only matrix $\mathbf{A}_k$ is unknown. Moreover, since $\mathbf{F}_k = \mathbf{N}_k\mathbf{A}_k$ and $\mathbf{N}_k$ is made up of orthonormal columns, one can see that the transmitted power constraint in (4.37) becomes

$$\mathrm{tr}\{\mathbf{A}_k\mathbf{A}_k^H\} = m_k. \tag{4.44}$$

Now, our task is to design processing matrices $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_K$, $\mathbf{D}_1, \mathbf{D}_2, ..., \mathbf{D}_K$, and $\mathbf{B}$. To solve this problem, we introduce the following lemma, which is proved in Appendix F.

Lemma 4.1: If the ZF criterion (4.38) is satisfied, the received signal vector at MS_k is

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{F}_k\mathbf{B}_{k,k}^{-1}\mathbf{v}_k + \mathbf{n}_k = \mathbf{H}_k\mathbf{N}_k\mathbf{A}_k\mathbf{B}_{k,k}^{-1}\mathbf{v}_k + \mathbf{n}_k . \tag{4.45}$$

After the linear receiver processing this received signal vector becomes

$$\mathbf{y}_k' = \mathbf{D}_k\mathbf{H}_k\mathbf{F}_k\mathbf{B}_{k,k}^{-1}\mathbf{v}_k + \mathbf{D}_k\mathbf{n}_k = \mathbf{D}_k\mathbf{H}_k\mathbf{N}_k\mathbf{A}_k\mathbf{B}_{k,k}^{-1}\mathbf{v}_k + \overline{\mathbf{n}}_k . \tag{4.46}$$

From Lemma 4.1 one can see that the nonlinear pre-processing in the structure shown in Figure 4.7 effectively decomposes the multi-user MIMO channel into parallel independent single-user MIMO channels as in the NL-DECOM technique. The data streams of MS_k can be seen as passing through an equivalent single-user MIMO channel $\mathbf{H}_k\mathbf{N}_k$ without any MUI. It is noted that the equivalent single-user MIMO channel $\mathbf{H}_k\mathbf{N}_k$ here is the same as that formed by the NL-DECOM technique. According to (4.46), matrices $\mathbf{D}_k$, $\mathbf{A}_k$, and $\mathbf{B}_{k,k}$ can be seen as the linear receiver-processing matrix, the feedforward filter at the transmitter, and the feedback filter at the transmitter for MS_k in the decomposed single-user MIMO channel. One can also see that the performance of MS_k is only influenced by processing matrices $\mathbf{A}_k$, $\mathbf{D}_k$, and $\mathbf{B}_{k,k}$, so processing matrices $\mathbf{A}_k$, $\mathbf{D}_k$, and $\mathbf{B}_{k,k}$ (k = 1,...,K) can be designed separately by optimizing the performance of each mobile. Next, we will discuss how to design these processing matrices for each mobile.

For MS_k, to satisfy the ZF criterion expressed by (4.39), clearly,

$$\mathbf{D}_k\mathbf{H}_k\mathbf{F}_k = \mathbf{D}_k\mathbf{H}_k\mathbf{N}_k\mathbf{A}_k = \mathbf{B}_{k,k} \tag{4.47}$$

is required, so from (4.46) $\mathbf{y}_k' = \mathbf{v}_k + \overline{\mathbf{n}}_k$. Therefore, for each data stream, all the spatial interference from other data streams is completely eliminated, and the performance of MS_k will only be determined by the noise $\overline{\mathbf{n}}_k$. The covariance matrix of $\overline{\mathbf{n}}_k$ is

$$\mathbf{\Phi}_{\bar{\mathbf{n}}_k\bar{\mathbf{n}}_k} \equiv \mathrm{E}[\bar{\mathbf{n}}_k\bar{\mathbf{n}}_k^H] = \sigma_n^2\mathbf{D}_k\mathbf{D}_k^H \tag{4.48}$$

and the variances of noises are:

$$(\sigma^2_{\bar{n}_k^{(1)}}, \sigma^2_{\bar{n}_k^{(2)}}, \ldots, \sigma^2_{\bar{n}_k^{(m_k)}})^T = \sigma_n^2 \mathrm{diag}(\mathbf{D}_k\mathbf{D}_k^H). \tag{4.49}$$

Since at high SNR the performance of MS_k is dominated by the noise with the largest variance, we intend to find the solution for $\mathbf{A}_k$, $\mathbf{D}_k$, and $\mathbf{B}_{k,k}$ which satisfies (4.47) and is optimal in the minimax noise variance sense, i.e., the solution that satisfies (4.50). This solution is optimal in the minimum BER sense at high SNR.

$$\begin{aligned}
&\min_{\mathbf{A}_k,\mathbf{D}_k,\mathbf{B}_{k,k}} \max\{\sigma^2_{\bar{n}_k^{(1)}}, \sigma^2_{\bar{n}_k^{(2)}}, \ldots, \sigma^2_{\bar{n}_k^{(m_k)}}\} \\
&\text{subject to} \quad (\sigma^2_{\bar{n}_k^{(1)}}, \sigma^2_{\bar{n}_k^{(2)}}, \ldots, \sigma^2_{\bar{n}_k^{(m_k)}})^T = \sigma_n^2 \mathrm{diag}(\mathbf{D}_k\mathbf{D}_k^H), \\
&\qquad\qquad \mathbf{D}_k\mathbf{H}_k\mathbf{N}_k\mathbf{A}_k = \mathbf{B}_{k,k}, \\
&\qquad\qquad \mathbf{B}_{k,k} \text{ is a unit upper triangular matrix}, \\
&\qquad\qquad \mathrm{tr}\{\mathbf{A}_k\mathbf{A}_k^H\} = m_k,
\end{aligned} \tag{4.50}$$

where the last equation is the transmitted power constraint for MS_k as in (4.44).

It is shown in Appendix G that the solution of (4.50) can be given by the following equations.

$$\begin{aligned}
&\mathbf{H}_k\mathbf{N}_k = \mathbf{Q}_k\mathbf{R}_k\mathbf{P}_k^H, \\
&\mathbf{G}_k = \mathrm{diag}[[\mathbf{R}_k]_{11}^{-1}, \ldots, [\mathbf{R}_k]_{m_k m_k}^{-1}], \\
&\mathbf{D}_k = \mathbf{G}_k\mathbf{Q}_k^H, \\
&\mathbf{B}_{k,k} = \mathbf{G}_k\mathbf{R}_k, \\
&\mathbf{A}_k = \mathbf{P}_k,
\end{aligned} \tag{4.51}$$

where $\mathbf{H}_k\mathbf{N}_k = \mathbf{Q}_k\mathbf{R}_k\mathbf{P}_k^H$ is the geometric mean decomposition (GMD) [90], [94] (or the equal-diagonal QR decomposition [89]) of $\mathbf{H}_k\mathbf{N}_k$, and $\mathbf{G}_k$ is a diagonal matrix that makes $\mathbf{B}_{k,k} = \mathbf{G}_k\mathbf{R}_k$ a unit triangular matrix.

The GMD of a matrix is defined in [89], [90], and [94]. If $\mathbf{X}$ is a rank K_r matrix of size $M \times N$, the GMD of $\mathbf{X}$ is $\mathbf{X} = \mathbf{QRP}^H$, where $\mathbf{P}$ (size $N \times K_r$) and $\mathbf{Q}$

(size $M \times K_r$) have orthonormal columns, and $\mathbf{R}$ is a $K_r \times K_r$ real upper triangular matrix with equal diagonal elements and

$$[\mathbf{R}]_{ii} = \bar{\sigma} \equiv (\prod_{\sigma_j > 0} \sigma_j)^{1/K_r}, \quad 1 \le i \le K_r . \tag{4.52}$$

Here, the σ_j ($j = 1,..., K_r$) are the singular values of $\mathbf{X}$ and $\bar{\sigma}$ is the geometric mean of the positive singular values of $\mathbf{X}$. The ways to perform the GMD of a matrix are described in [89] and [94].

Since it is assumed that the entries of $\mathbf{H}_k$ are normalized i.i.d. zero-mean complex Gaussian random variables and $\mathbf{N}_k$ is found to be an $N \times (N - \sum_{i=k+1}^{K} m_i)$ matrix with orthonormal columns, the entries of the decomposed single-user MIMO channel $\mathbf{H}_k\mathbf{N}_k$ (size $m_k \times (N - \sum_{i=k+1}^{K} m_i)$) formed in Lemma 4.1 are still normalized i.i.d. zero-mean complex Gaussian random variables, so the rank of $\mathbf{H}_k\mathbf{N}_k$ is m_k with probability one. Therefore, in the solution of (4.51) $\mathbf{D}_k$ and $\mathbf{B}_{k,k}$ are of size $m_k \times m_k$, while $\mathbf{A}_k$ is of size $(N - \sum_{i=k+1}^{K} m_i) \times m_k$. Assuming $\sigma_k^{(j)}$ ($j = 1,..., m_k$) are the singular values of $\mathbf{H}_k\mathbf{N}_k$, from (4.51) and (4.52) it can be found that

$$\mathbf{G}_k = \bar{\sigma}_k^{-1} \mathbf{I}_{m_k \times m_k} \tag{4.53}$$

where $\bar{\sigma}_k \equiv (\prod_{j=1}^{m_k} \sigma_k^{(j)})^{1/m_k}$ is the geometric mean of the singular values of $\mathbf{H}_k\mathbf{N}_k$. From (4.49), (4.51), and (4.53), one can find that the variances of $\bar{n}_k^{(1)}, \bar{n}_k^{(2)}, ..., \bar{n}_k^{(m_k)}$ are all equal, i.e.,

$$\sigma^2_{\bar{n}_k^{(1)}} = \sigma^2_{\bar{n}_k^{(2)}} = ... = \sigma^2_{\bar{n}_k^{(m_k)}} = \sigma_n^2 \bar{\sigma}_k^{-2} . \tag{4.54}$$

We use one symbol

$$\sigma^2_{\bar{n}_k} \equiv \sigma_n^2 \bar{\sigma}_k^{-2} \tag{4.55}$$

to represent this value. Clearly, if all $\mathbf{A}_k$, $\mathbf{D}_k$, and $\mathbf{B}_{k,k}$ (k = 1,...,K) are found through (4.51), all the receive branches of one mobile will have the same performance.

Now, let us summarize how to find the solution for processing matrices **F**, **D**, and **B** in the proposed ZF nonlinear joint Tx-Rx processing algorithm. First, $\mathbf{N}_k$ (k = 1,...,K) should be determined using (4.43). Then, (4.51) is used to find processing matrices $\mathbf{A}_k$ and $\mathbf{D}_k$ (k = 1,...,K). When all $\mathbf{N}_k$, $\mathbf{A}_k$, and $\mathbf{D}_k$ (k = 1,...,K) are found, clearly, processing matrices $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_K]$ where $\mathbf{F}_k = \mathbf{N}_k \mathbf{A}_k$ and **D** (see (4.35)) are determined. Finally, the processing matrix **B** can be determined through (4.38), i.e., **B** = **DHF**.

Finally, we'd like to point out that it can be shown that processing matrix **F** found in the ZF nonlinear joint Tx-Rx processing algorithm satisfies $\mathbf{F}^H \mathbf{F} = \mathbf{I}$, so the average transmitted energy per symbol interval for the ith data stream is $E[[\mathbf{F}]_i |\tilde{a}_i|^2 [\mathbf{F}]_i^H] = \sigma_{\tilde{a}}^2$ (i = 1,...,M), where $\tilde{a}_i$ denotes the ith element of $\tilde{\mathbf{a}}$. One can see that the transmitted power is allocated equally to all the data streams.

To show the advantage of the proposed ZF nonlinear joint Tx-Rx processing algorithm, next, we will compare it with the conventional ZF nonlinear pre-processing algorithm described in Section 2.3. It is noted that in the conventional ZF nonlinear pre-processing algorithm the transmitted power is also allocated equally to all the data streams.

4.3.2.1 Advantage of the ZF Nonlinear Joint Tx-Rx Processing Algorithm over the Conventional ZF Nonlinear Pre-Processing Algorithm

The advantage of the ZF nonlinear joint Tx-Rx processing algorithm proposed in this section over the conventional ZF nonlinear pre-processing algorithm described in Section 2.3 can be shown by comparing the diversity order of these two algorithms.

When the proposed ZF nonlinear joint Tx-Rx processing algorithm is applied, from $\mathbf{y}_k' = \mathbf{v}_k + \bar{\mathbf{n}}_k$ and (4.54) one can see that the performance of MS_k is decided by the geometric mean of the singular values of $\mathbf{H}_k \mathbf{N}_k$. Since we have revealed

that $\mathbf{H}_k\mathbf{N}_k$ is an $m_k \times (N - \sum_{i=k+1}^{K} m_i)$ matrix whose entries are normalized i.i.d. zero-mean complex Gaussian random variables, using the result in [95], it can be found the diversity order of MS_k is

$$d_{MS_k} = m_k((N - \sum_{i=k+1}^{K} m_i) + 1 - m_k) = m_k(N - \sum_{i=k}^{K} m_i + 1). \quad (4.56)$$

Since the overall system performance is dominated by the performance of the worst mobile, the diversity order of the system is

$$d = \min\{m_k(N - \sum_{i=k}^{K} m_i + 1) : 1 \le k \le K\}. \quad (4.57)$$

Using the same analysis approach as that in [96], it can be found that the diversity order of the conventional ZF nonlinear pre-processing algorithm described in Section 2.3 is $(N - M + 1)$. Since

$$\min\{m_k(N - \sum_{i=k}^{K} m_i + 1) : 1 \le k \le K\} > (N - M + 1) \quad (4.58)$$

always holds for multi-user MIMO systems with multiple-antenna mobiles, one can see that higher system diversity order can be achieved by using the ZF nonlinear joint Tx-Rx processing algorithm proposed in this section, which implies the proposed algorithm can achieve better performance than the conventional ZF nonlinear pre-processing algorithm at high SNR.

Simulation results in Section 4.3.2.4 will show that the ZF nonlinear joint Tx-Rx processing algorithm we have proposed in this section can provide significant performance improvement as compared to the conventional ZF nonlinear pre-processing algorithm.

4.3.2.2 Maximum Achievable Sum Rate in a System Employing the ZF Nonlinear Joint Tx-Rx Processing Algorithm

Using the same analysis as in [76] and [84], we can find that the maximum achievable sum rate in a system employing the ZF nonlinear joint Tx-Rx processing algorithm is

$$R_{ZF} = \max_{\{p_k^{(i)}\}} \sum_{k=1}^{K} \sum_{i=1}^{m_k} \Re(\sqrt{3p_k^{(i)}/2}, \sigma_n / \bar{\sigma}_k) \quad (4.59)$$

$$\text{subject to } \sum_{k=1}^{K}\sum_{i=1}^{m_k} p_k^{(i)} = M$$

where function $\Re(\cdot,\cdot)$ is defined in (4.26), $\bar{\sigma}_k$ is the geometric mean of the singular values of the decomposed single-user MIMO channel for MS_k, $p_k^{(i)}$ represents the power allocated to the ith data stream of MS_k, and the constraint $\sum_{k=1}^{K}\sum_{i=1}^{m_k} p_k^{(i)} = M$ is applied to keep the total transmitted power constant. Since finding the maximum achievable sum rate from (4.59) is in general a non-convex problem, the power allocation factors $p_k^{(i)}$ that achieve the maximum achievable sum rate are not easy to find in closed-form. In this thesis, we only consider the maximum achievable sum rate when the SNR is high. In this case, the transmitted power is allocated almost equally to all data streams [76], i.e., $p_k^{(i)} = 1$. With the high SNR assumption, using the same approach as that in [76] and [84] it can be found that the maximum achievable sum rate is

$$R_{ZF} = \sum_{k=1}^{K} m_k \log_2(\bar{\sigma}_k^2 / \sigma_n^2) - M \log_2(\pi e / 6). \qquad (4.60)$$

Moreover, it can be shown that (4.60) is also the maximum achievable sum rate of the decomposed multiple single-user MIMO channels $\mathbf{H}_1\mathbf{N}_1, \mathbf{H}_2\mathbf{N}_2, ..., \mathbf{H}_K\mathbf{N}_K$ at high SNR. In other words, the ZF nonlinear joint Tx-Rx processing algorithm achieves the maximum achievable sum rate of the decomposed multiple single-user MIMO channels $\mathbf{H}_1\mathbf{N}_1, \mathbf{H}_2\mathbf{N}_2, ..., \mathbf{H}_K\mathbf{N}_K$ obtained from Lemma 4.1 at high SNR.

4.3.2.3 Ordering in the ZF Nonlinear Joint Tx-Rx Processing Algorithm

We have shown that the "best-first" ordering method can achieve the optimal order in the minimax noise variance sense for the conventional ZF nonlinear pre-processing algorithm in Section 3.4. Ordering should be performed on the rows of the channel matrix $\mathbf{H}$. For the ZF nonlinear joint Tx-Rx processing algorithm, due to the receiver processing applied at each mobile the rows of the channel matrix $\mathbf{H}$ cannot be ordered independently. However, the channel matrices of different mobiles, i.e., $\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_K$ in $\mathbf{H}$, can be ordered to improve the system

performance. In this section, we will first describe the change of system structure caused by ordering the channel matrices of different mobiles and then propose a combined optimal diversity and "best-first" (CODBF) ordering method.

4.3.2.3.1 Change of System Structure Due to Ordering

The system structure when no ordering is applied has been shown in Figure 4.7. In this section, we use $L \equiv \{l_1, l_2, ..., l_K\}$ to represent an order of the channel matrices of different mobiles, and it should be one permutation of the mobile indices, 1,2,…,K. If the order of $\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_K$ is found to be $L \equiv \{l_1, l_2, ..., l_K\}$, $\mathbf{H}$ should be ordered into

$$\mathbf{H}^{(L)} \equiv [\mathbf{H}_{l_1}^T, \mathbf{H}_{l_2}^T, ..., \mathbf{H}_{l_K}^T]^T. \tag{4.61}$$

Then, processing matrices $\mathbf{F}$, $\mathbf{D}$, and $\mathbf{B}$ should be generated using $\mathbf{H}^{(L)}$ instead of $\mathbf{H}$, and they are denoted as $\mathbf{F}^{(L)}$, $\mathbf{D}^{(L)}$, and $\mathbf{B}^{(L)}$. The system structure when order L is used is shown in Figure 4.8. Now, the received signal vector that will go into the modulo and decision devices is

$$\begin{aligned} \mathbf{y}^{(L)\prime} &= \mathbf{D}^{(L)}\mathbf{H}^{(L)}\mathbf{x}^{(L)} + \mathbf{D}^{(L)}\mathbf{n}^{(L)} \\ &= \mathbf{D}^{(L)}\mathbf{H}^{(L)}\mathbf{F}^{(L)}\mathbf{B}^{(L)-1}\mathbf{v}^{(L)} + \mathbf{D}^{(L)}\mathbf{n}^{(L)}, \\ &= \mathbf{v}^{(L)} + \mathbf{D}^{(L)}\mathbf{n}^{(L)} \end{aligned} \tag{4.62}$$

where $\mathbf{y}^{(L)\prime} \equiv (\mathbf{y}_{l_1}', \mathbf{y}_{l_2}', ..., \mathbf{y}_{l_K}')^T$ is made up of the received signal vectors at $\mathrm{MS}_{l_1}, \mathrm{MS}_{l_2}, \ldots, \mathrm{MS}_{l_K}$ and $\mathbf{n}^{(L)} \equiv (\mathbf{n}_{l_1}, \mathbf{n}_{l_2}, ..., \mathbf{n}_{l_K})^T$ is made up of the noise vectors at $\mathrm{MS}_{l_1}, \mathrm{MS}_{l_2}, \ldots, \mathrm{MS}_{l_K}$. In this ordered system, to ensure the transmission of data vector $\mathbf{a}_k$ to MS_k, the data vector applied to the precoding device should also be ordered into $\mathbf{a}^{(L)} = (\mathbf{a}_{l_1}, \mathbf{a}_{l_2}, ..., \mathbf{a}_{l_K})^T$.

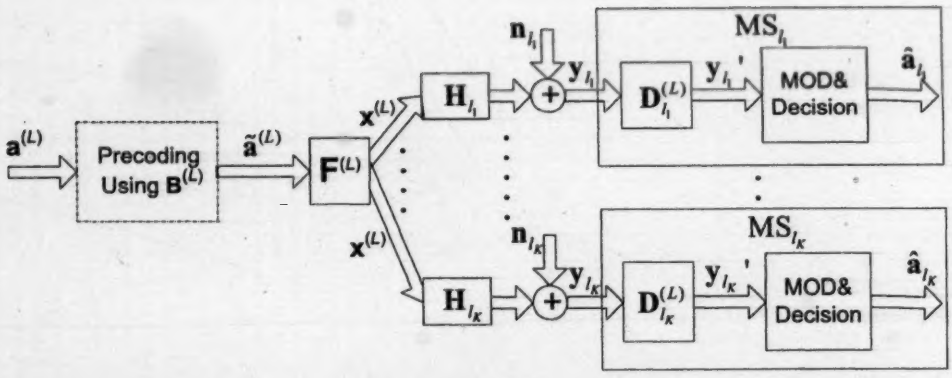


Figure 4.8 Structure of the ZF nonlinear joint Tx-Rx processing algorithm when order L is used.

4.3.2.3.2 Combined Optimal Diversity and "Best-First" Ordering Method

The CODBF ordering method we propose intends to improve the overall system performance for the ZF nonlinear joint Tx-Rx processing algorithm. This method has two steps described as follows. In the first step, it achieves the optimal order in the maximal diversity order sense, which means the system diversity order when the optimal order is used must be higher than or the same as the system diversity order when any other order is used. Then, based on the result of the first step, the "best-first" ordering method is applied to further improve the performance.

Step 1: Optimal Diversity Ordering

As we have known, without ordering the diversity order of MS_k is (4.56), and the diversity order of the system is (4.57). If order $L \equiv \{l_1, l_2, ..., l_K\}$ is applied, it can be found that the diversity order of MS_{l_k} is

$$d_{MS_{l_k}} = m_{l_k}(N - \sum_{i=k}^{K} m_{l_i} + 1), \tag{4.63}$$

and the diversity order of the system is

$$d^{(L)} = \min\{m_{l_k}(N - \sum_{i=k}^{K} m_{l_i} + 1) : 1 \le k \le K\}. \tag{4.64}$$

Clearly, for this system, the order which is optimal in the maximal diversity order sense, is preferred. In other words, we'd like to find the order $\bar{L}$ which

ensures $d^{(\overline{L})} \geq d^{(Q)}$ where Q represents any other order, and $d^{(\overline{L})}$ and $d^{(Q)}$ represent the diversity orders of the system when order $\overline{L}$ and Q are used, respectively. It is shown in Appendix H that order $\overline{L} \equiv \{\overline{l}_1, \overline{l}_2, \ldots, \overline{l}_K\}$ which satisfies

$$m_{\overline{l}_i} \geq m_{\overline{l}_j} \qquad \text{if } i < j \tag{4.65}$$

is the optimal order in the maximal diversity order sense. Here, $m_{\overline{l}_i}$ and $m_{\overline{l}_j}$ are the numbers of antennas at $\text{MS}_{\overline{l}_i}$ and $\text{MS}_{\overline{l}_j}$, respectively. This order can be found by using the method described below.

$\overline{L}$ is formed from $\overline{l}_1, \overline{l}_2$, to $\overline{l}_K$. To find $\overline{l}_i$, all the mobiles whose indices have not been ordered into $\overline{l}_1, \ldots, \overline{l}_{i-1}$ will be searched, and the index of the mobile with the largest number of antennas will be chosen as $\overline{l}_i$. When there are more than one mobile with the largest number of antennas, any one of them can be ordered into the position $\overline{l}_i$.

Example 4.3 below is used to show how the optimal diversity ordering method works.

Example 4.3. Let us consider a multi-user MIMO system, in which there are 7 transmit antennas at the base station and 3 mobiles. The numbers of antennas at the mobiles are $m_1 = 2$ for MS_1, $m_2 = 3$ for MS_2, and $m_3 = 2$ for MS_3. Without ordering the system diversity order can be found from (4.57), and it is $d = 2$. If the optimal diversity ordering proposed above is used, the order of the channel matrices of different mobiles should be $\overline{S} \equiv \{\overline{s}_1, \overline{s}_2, \overline{s}_3\} = \{2,1,3\}$ or $\overline{S} = \{2,3,1\}$. With this order $m_{\overline{s}_1} = 3$ and $m_{\overline{s}_2} = m_{\overline{s}_3} = 2$, so one can see that (4.65) is satisfied. The system diversity order when this order is used can be found from (4.64), and it is increased to $d^{(\overline{S})} = 3$.

It is noted that the above optimal diversity ordering only works for the systems, in which the mobiles have different numbers of antennas, since from (4.64) one can see that ordering the channel matrices of different mobiles will not

change the system diversity order when all the mobiles have the same number of antennas.

Step 2: "Best-First" Ordering

It can be seen that the optimal diversity ordering in Step 1 does not provide an ordering rule for the mobiles with the same number of antennas. In the $\overline{L}$ found in Step 1, the indices of the mobiles who have equal number of antennas are adjacent, but they are ordered randomly. (For example, in the Example 4.3 given above the indices of MS_1 and MS_3, 1 and 3, are ordered randomly as $\overline{s}_2$ and $\overline{s}_3$.) These indices can also be ordered properly to achieve better performance.

Assume $\overline{\mathbf{l}}_i \equiv \{\overline{l}_i, \overline{l}_{i+1}, ..., \overline{l}_{i+p}\}$ is a subset of $\overline{L}$ which is made up of the indices of a group of mobiles with equal number of antennas. From (4.54) and (4.55), it is seen that the performance of $MS_{\overline{l}_j}$ is determined by the noise variance $\sigma^2_{\overline{n}_{\overline{l}_j}}$. Therefore, the overall performance of this group of mobiles will be dominated by the largest noise variance, i.e., $\max\{\sigma^2_{\overline{n}_{\overline{l}_{i+t}}} : 1 \le t \le p\}$, and hence we intend to find the optimal order in the minimax noise variance sense. That means, if $\mathbf{l}_i \equiv \{l_i, l_{i+1}, ..., l_{i+p}\}$ is the optimal order found for $\overline{\mathbf{l}}_i$, $\max\{\sigma^2_{\overline{n}_{\overline{l}_{i+t}}} : 1 \le t \le p\}$ must be smaller than or the same as the maximal noise variance when any other order is used. It is shown in Appendix I that the "best-first" ordering method can achieve the optimal order in the minimax noise variance sense for the group of mobiles with equal number of antennas. A detailed explanation of the working mechanism of the "best-first" ordering can be found in Section 3.3. In the following a brief description of the "best-first" ordering for the ZF nonlinear joint Tx-Rx processing algorithm is given.

Assume $\mathbf{l}_i \equiv \{l_i, l_{i+1}, ..., l_{i+p}\}$ is the final order found for $\overline{\mathbf{l}}_i$. Using the "best-first" ordering method, $l_i, l_{i+1}, ..., l_{i+p}$ should be found from l_i, l_{i+1}, to l_{i+p} iteratively. To find l_{i+t}, all the unordered mobiles (the mobiles whose indices have not been chosen as $l_i, ..., l_{i+t-1}$) are tested. The index of the mobile that will lead to the minimum noise variance $\sigma^2_{\overline{n}_{\overline{l}_{i+t}}}$ will be ordered into the position l_{i+t}.

All the groups of mobiles with equal number of antennas should be ordered in this manner. It is noted that when all the mobiles have the same number of antennas, the optimal diversity ordering in Step 1 does not work, so the "best-first" ordering method described here is used directly.

Simulation results in the next section will show that the CODBF ordering method proposed in this section can greatly improve the performance of the ZF nonlinear joint Tx-Rx processing algorithm.

4.3.2.4 Simulation Results

In this section, simulations are used to show the advantages of the ZF nonlinear joint Tx-Rx processing algorithm and the CODBF ordering method proposed in Section 4.3.2.3.

In our simulation, the flat Rayleigh fading channel is assumed to change independently from one symbol interval to the next. The SNR is defined by (2.43). The BER is the average BER of all the mobiles.

Figure 4.9 presents BER performance comparison of the proposed ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm. A multi-user MIMO system, in which there are 6 transmit antennas at the base station and 2 mobiles each with 3 antennas, i.e., $N = 6$, $K = 2$, and $m_1 = m_2 = 3$, is considered. QPSK with Gray encoding [74] is used. The performance of the ZF nonlinear joint Tx-Rx processing algorithm without ordering and with the ordering described in Section 4.3.2.3 is shown. It is noted that since all the mobiles have the same number of antennas, in the ordered case, only the "best-first" ordering described in Section 4.3.2.3.2 Step 2 is used. For the conventional ZF nonlinear pre-processing algorithm, performance curves of this algorithm without ordering and with the "best-first" ordering described in Section 3.4 are drawn. It is clearly seen that the proposed ZF nonlinear joint Tx-Rx processing algorithm can achieve significantly better performance than the conventional ZF nonlinear pre-processing algorithm. At BER = 10^{-3} the ZF nonlinear joint Tx-Rx processing algorithm without ordering achieves approximately 10 dB gain over the conventional ZF nonlinear pre-processing algorithm without ordering, and the ZF nonlinear joint Tx-Rx processing

algorithm with ordering achieves approximately 4 dB gain over the conventional ZF nonlinear pre-processing algorithm with ordering. From this figure, it can also be found that the diversity order of the ZF nonlinear joint Tx-Rx processing algorithm, which is 3 (see (4.57)), is much higher than the diversity order of the conventional ZF nonlinear pre-processing algorithm, which is only 1 (= N-M+1). The improved performance attained by ordering is also shown in this figure. At BER = 10^{-3}, using the "best-first" ordering, approximately 1 dB gain and 7 dB gain are obtained for the ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm, respectively.

Figure 4.10 shows performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm in the downlink of a multi-user MIMO system with 8 transmit antennas at the base station and 3 mobiles (N = 8 and K = 3). It is assumed that there are 2 antennas at MS_1, 3 antennas at MS_2 and MS_3 ($m_1 = 2$ and $m_2 = m_3 = 3$). QPSK with Gray encoding [74] is used. It can be seen that at BER = 10^{-3} the ZF nonlinear joint Tx-Rx processing algorithm without ordering achieves about 6 dB gain over the conventional ZF nonlinear pre-processing algorithm without ordering. For the ordered case, the "best-first" ordering described in Section 3.4 is used for the conventional ZF nonlinear pre-processing algorithm and the CODBF ordering method described in Section 4.3.2.3.2 is used for the ZF nonlinear joint Tx-Rx processing algorithm. It can be found that although the ordered ZF nonlinear joint Tx-Rx processing algorithm is a little worse than the ordered conventional ZF nonlinear pre-processing algorithm at low SNR, since the diversity order of the former, which is 3 (see (4.64)), is larger than that of the latter, which is 1 (= N-M+1), the advantage of the ordered ZF nonlinear joint Tx-Rx processing algorithm appears at high SNR. At BER = 10^{-3}, the ZF nonlinear joint Tx-Rx processing algorithm with ordering achieves about 2 dB gain over the conventional ZF nonlinear pre-processing algorithm with ordering. Using the CODBF or "best-first" ordering, approximately 4 dB or 8 dB gain is attained for the ZF nonlinear joint Tx-Rx processing algorithm or the conventional ZF nonlinear pre-processing algorithm at BER = 10^{-3}. Also, it can be found the

CODBF ordering method increases the diversity order of the ZF nonlinear joint Tx-Rx processing algorithm from 2 (see (4.57)) to 3 (see (4.64)).

In Figure 4.11, we compare the BER performance of the ZF nonlinear joint Tx-Rx processing algorithm and the L-DECOM technique combined with different single-user MIMO processing algorithms [59], [60]. A multi-user MIMO system, in which there are 6 transmit antennas at the base station and 3 mobiles each with 2 antennas ($N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$), is considered. QPSK with Gray encoding [74] is used. In our simulation, the ZF linear detector (LD) [30], [44], the ZF DFD with the "best-first" ordering [18], and the MLD [30]–[32] are used as the single-user MIMO processing algorithms when the L-DECOM technique is used, and their curves are named as L-DECOM (ZF LD), L-DECOM (ZF DFD), and L-DECOM (MLD), respectively. It is noted that the L-DECOM (ZF LD) is a linear joint Tx-Rx processing algorithm, while the L-DECOM (ZF DFD) and L-DECOM (MLD) are nonlinear joint Tx-Rx processing algorithms, which are made up of linear transmitter pre-processing and nonlinear receiver processing. The ordering method described in Section 4.3.2.3.2 is used for the ZF nonlinear joint Tx-Rx processing algorithm. From this figure, it is clearly seen that the ZF nonlinear joint Tx-Rx processing algorithm can achieve significantly better performance than the other algorithms. One should note that the ZF nonlinear joint Tx-Rx processing algorithm even outperforms the L-DECOM (MLD) algorithm, which has much higher computational complexity. At $\mathrm{BER} = 10^{-3}$ the ZF nonlinear joint Tx-Rx processing algorithm achieves approximately 13 dB gain over the L-DECOM (ZF LD) algorithm, 9.5 dB gain over the L-DECOM (ZF DFD) algorithm, and 1.5 dB gain over the L-DECOM (MLD) algorithm.

Finally, we compare the BER performance of the ZF nonlinear joint Tx-Rx processing algorithm and the NL-DECOM technique combined with different single-user MIMO processing algorithms in Figure 4.12. A multi-user MIMO system, in which there are 9 transmit antennas at the base station and 3 mobiles each with 3 antennas ($N = 9$, $K = 3$, and $m_1 = m_2 = m_3 = 3$), is considered. 16QAM with Gray encoding [74] is used. The ZF linear pre-processing (LP) algorithm

[40], [43] and the precoded ZF DFD algorithm [89], [90] are used as the self-interference cancellation algorithms when the NL-DECOM technique is applied, and their curves are named as NL-DECOM (ZF LP) and NL-DECOM (precoded ZF DFD), respectively. The ordering method described in Section 4.3.2.3.2 is used for the ZF nonlinear joint Tx-Rx processing algorithm. For the NL-DECOM technique, the "best-first" ordering method proposed in Section 4.2.4 is applied. The performance advantage of the ZF nonlinear joint Tx-Rx processing algorithm over the other two algorithms can be seen clearly from this figure. One can see that at BER = 10^{-3} the ZF nonlinear joint Tx-Rx processing algorithm achieves approximately 11.5 dB gain over the NL-DECOM (ZF LP) algorithm and 1 dB gain over the NL-DECOM (precoded ZF DFD) algorithm. We know that in the ZF nonlinear joint Tx-Rx processing algorithm the nonlinear pre-processing pre-eliminates both the MUI and the self-interference. In the NL-DECOM (ZF LP) algorithm, the nonlinear pre-processing only pre-eliminates the MUI, and the self-interference is pre-equalized by linear pre-processing, which makes the performance of this algorithm worse than that of the ZF nonlinear joint Tx-Rx processing algorithm. The NL-DECOM (precoded ZF DFD) algorithm uses nonlinear pre-processing to pre-eliminate the MUI and nonlinear receiver processing to equalize the self-interference. Its performance is worse than that of the ZF nonlinear joint Tx-Rx processing algorithm since the error propagation inherent to the DFD. Moreover, the NL-DECOM (precoded ZF DFD) algorithm will lead to higher computional complexity of the mobiles compared to the ZF nonlinear joint Tx-Rx processing algorithm, since nonlinear receiver processing is applied in the former, while linear receiver processing is applied in the later.

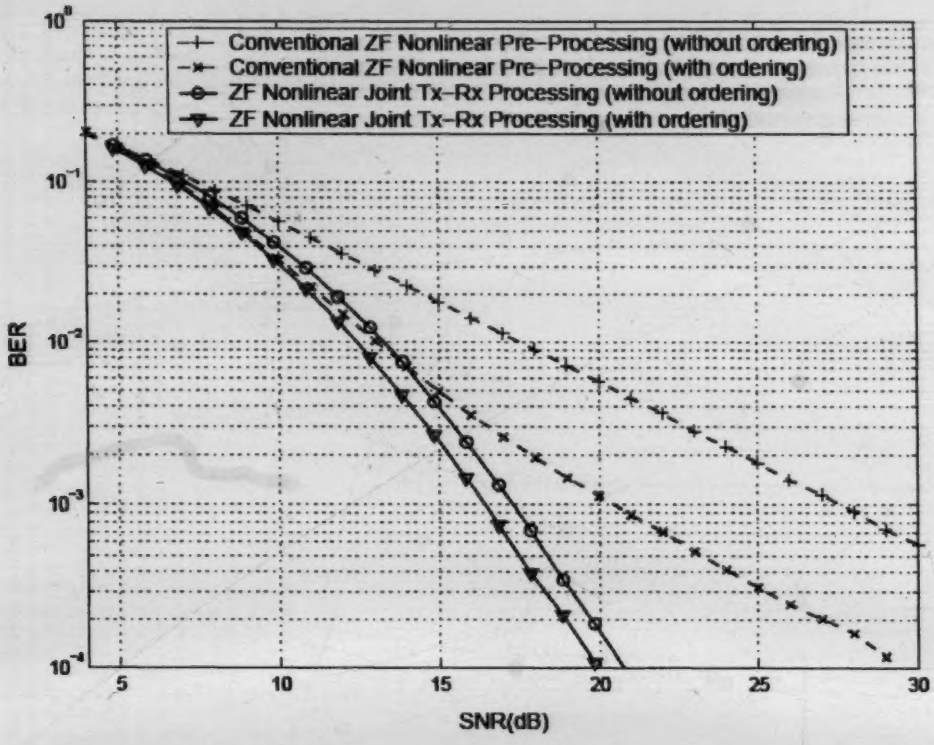


Figure 4.9 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $N = 6, K = 2$, and $m_1 = m_2 = 3$).

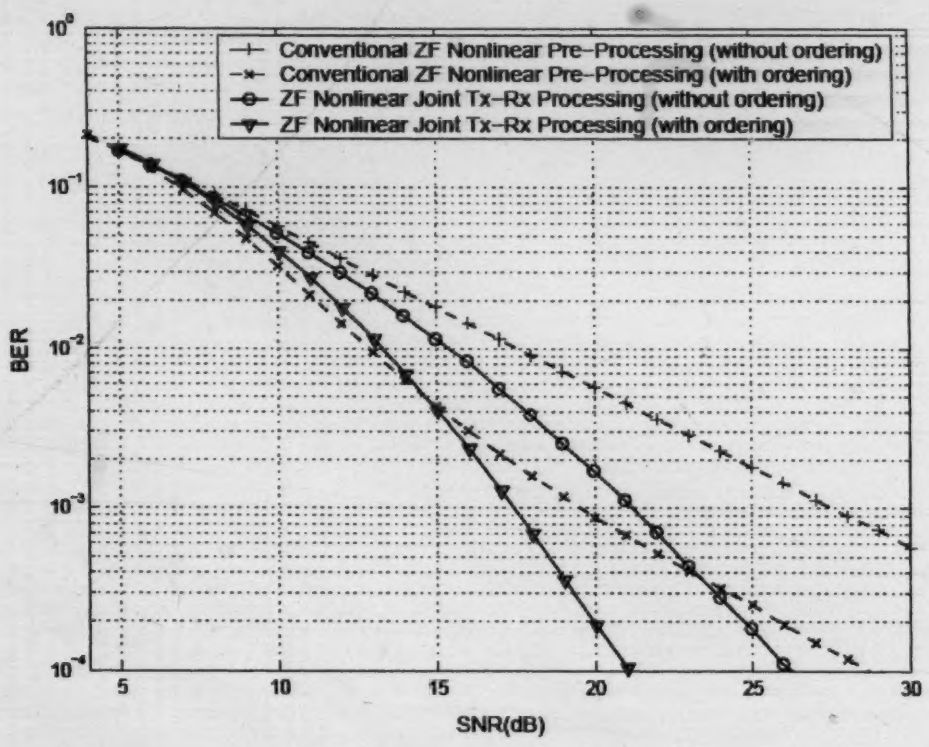


Figure 4.10 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the conventional ZF nonlinear pre-processing algorithm (QPSK, $N=8$, $K=3$, $m_1=2$, and $m_2=m_3=3$).

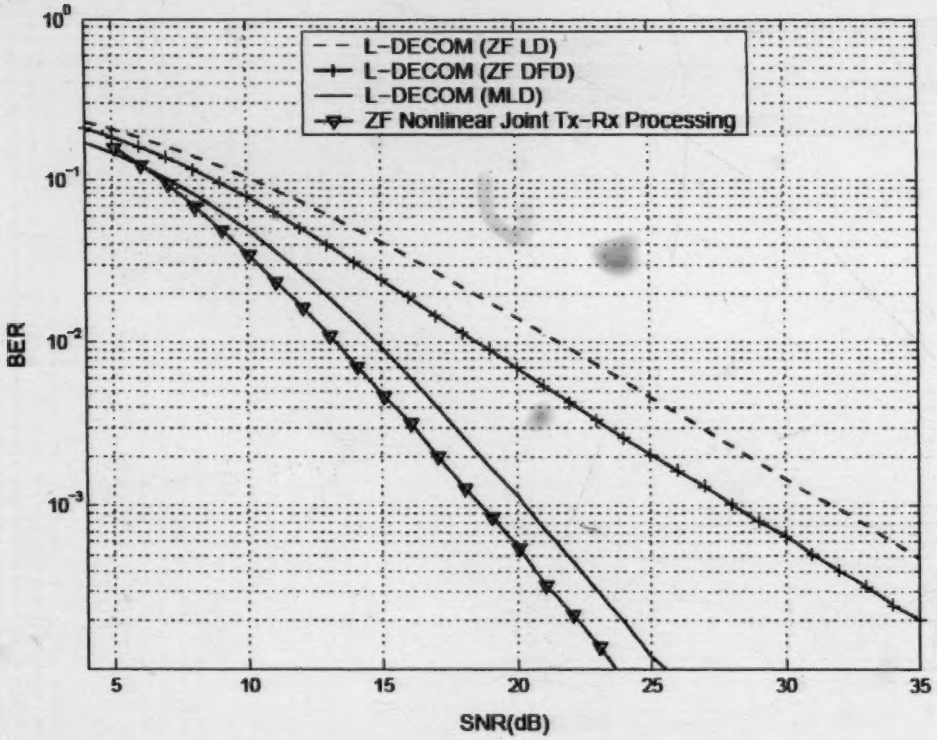


Figure 4.11 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the L-DECOM technique based algorithms (QPSK, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$).

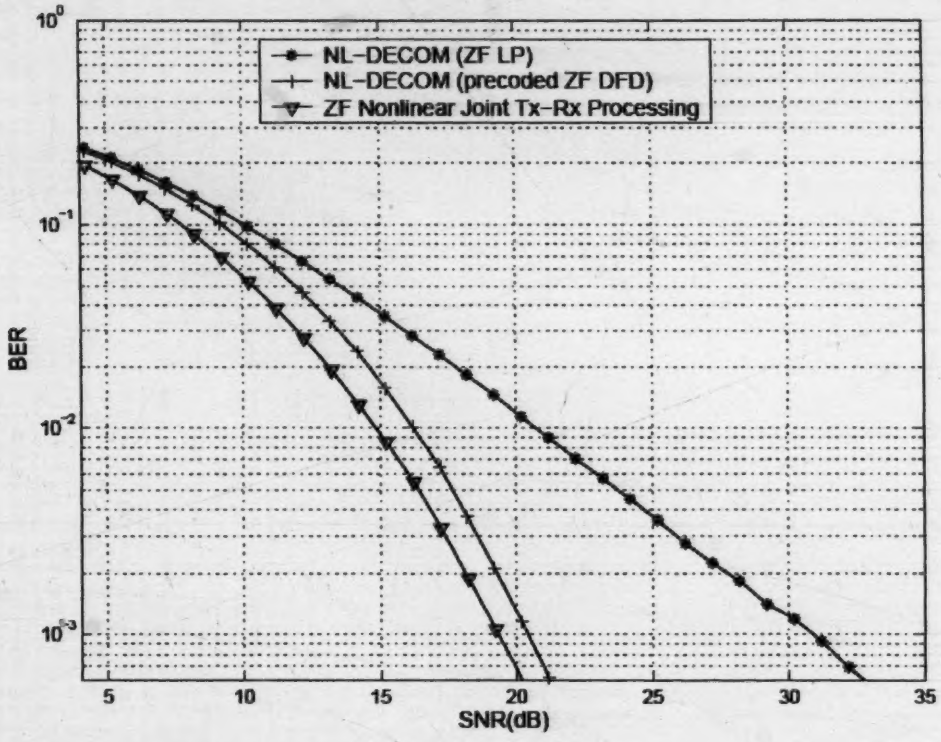


Figure 4.12 Performance comparison of the ZF nonlinear joint Tx-Rx processing algorithm and the NL-DECOM technique based algorithms (16QAM, $N = 9$, $K = 3$, and $m_1 = m_2 = m_3 = 3$).

4.3.2.5 Additional Comments

Although the ZF nonlinear joint Tx-Rx processing algorithm proposed in this section is designed for multi-user MIMO system with multiple-antenna mobiles, it can be found that this algorithm and the corresponding CODBF ordering method also work for the generalized multi-user MIMO system, in which some or even all mobiles have only one antenna. It can be found that in the generalized multi-user MIMO system any mobile with more than one antenna will have better performance using the ZF nonlinear joint Tx-Rx processing algorithm than using the conventional ZF nonlinear pre-processing algorithm. When used in the multi-user MIMO system with single-antenna mobiles, the proposed ZF nonlinear joint Tx-Rx processing algorithm is equivalent to the conventional ZF nonlinear pre-

processing algorithm, and the CODBF ordering method becomes the "best-first" ordering method. Therefore, the conventional ZF nonlinear pre-processing can be seen as a special case of the ZF nonlinear joint Tx-Rx processing algorithm.

4.3.3 MMSE Nonlinear Joint Tx-Rx Processing Algorithm

In this section, the MMSE criterion is used to design the processing matrices for the structure shown in Figure 4.7. It is assumed that the transmitter has the knowledge of the noise variance σ_n^2.

First, one can see that the error vector of this system is

$$\mathbf{e} = \mathbf{y}' - \mathbf{v} = \mathbf{Dy} - \mathbf{B}\tilde{\mathbf{a}} = (\mathbf{DHF} - \mathbf{B})\tilde{\mathbf{a}} + \mathbf{Dn}, \tag{4.66}$$

where

$$\mathbf{e} \equiv (\mathbf{e}_1^T, \mathbf{e}_2^T, ..., \mathbf{e}_K^T)^T, \tag{4.67}$$

and

$$\begin{aligned} \mathbf{e}_k &\equiv (e_k^{(1)}, e_k^{(2)}, ..., e_k^{(m_k)})^T \\ &= \mathbf{y}_k' - \mathbf{v}_k = \mathbf{D}_k\mathbf{y}_k - \mathbf{B}_k\tilde{\mathbf{a}} = (\mathbf{D}_k\mathbf{H}_k\mathbf{F} - \mathbf{B}_k)\tilde{\mathbf{a}} + \mathbf{D}_k\mathbf{n}_k \end{aligned} \tag{4.68}$$

is the error vector of MS_k. The error covariance matrices of $\mathbf{e}$ and $\mathbf{e}_k$ are

$$\mathbf{\Phi}_{\mathbf{ee}} \equiv \mathrm{E}[\mathbf{ee}^H] = (\mathbf{DHF} - \mathbf{B})\mathbf{R}_{\tilde{\mathbf{a}}}(\mathbf{DHF} - \mathbf{B})^H + \sigma_n^2\mathbf{DD}^H \tag{4.69}$$

and

$$\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k} \equiv \mathrm{E}[\mathbf{e}_k\mathbf{e}_k^H] = (\mathbf{D}_k\mathbf{H}_k\mathbf{F} - \mathbf{B}_k)\mathbf{R}_{\tilde{\mathbf{a}}}(\mathbf{D}_k\mathbf{H}_k\mathbf{F} - \mathbf{B}_k)^H + \sigma_n^2\mathbf{D}_k\mathbf{D}_k^H. \tag{4.70}$$

The MMSE-criterion-based solution of this system should be able to minimize $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{ee}}) = \sum_{k=1}^{K} \mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ with respect to processing matrices **B**, **F**, and **D**, where **B** is a unit upper triangular matrix and **D** is a block diagonal matrix. Meanwhile, constraint (4.37) is applied to constrain the transmitted power.

From (4.70), it is clear that $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ is only influenced by $\mathbf{D}_k$, while $\mathbf{D}_j$ ($j \neq k$) have no influence on $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$. Therefore, $\mathbf{D}_1$, $\mathbf{D}_2$, ..., $\mathbf{D}_K$ that minimize $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{ee}})$ can be found by minimizing $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_1\mathbf{e}_1})$, $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_2\mathbf{e}_2})$, ..., $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_K\mathbf{e}_K})$ separately. $\mathbf{D}_k$ that minimizes $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ can be found by differentiation of $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ with respect to $\mathbf{D}_k$ and setting the result to 0. As a result,

$$\mathbf{D}_k = \mathbf{B}_k \mathbf{F}^H \mathbf{H}_k^H (\mathbf{H}_k \mathbf{F}\mathbf{F}^H \mathbf{H}_k^H + \xi \mathbf{I})^{-1} \tag{4.71}$$

where $\xi = \sigma_n^2 / \sigma_a^2$. With $\mathbf{D}_k$ found in (4.71), one can then obtain

$$\mathbf{\Phi}_{\mathbf{e}_k \mathbf{e}_k} = \sigma_n^2 \mathbf{B}_k (\mathbf{F}^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{F} + \xi \mathbf{I})^{-1} \mathbf{B}_k^H . \tag{4.72}$$

It is noted that $\mathbf{D}_k$ found in (4.71) and $\mathbf{\Phi}_{\mathbf{e}_k \mathbf{e}_k}$ in (4.72) are functions of $\mathbf{B}_k$ and $\mathbf{F}$. Next, matrices $\mathbf{B}$ and $\mathbf{F}$ should be found by minimizing $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{ee}})$. However, a closed form solution is hard to find. To find a closed form solution, constraint on the processing matrix $\mathbf{F}$ is applied as follows.

We assume the matrix $\mathbf{F}$ can make $\mathbf{HF}$ a block upper triangular matrix, i.e.,

$$\mathbf{H}_j \mathbf{F}_k = \mathbf{0}_{m_j, m_k}, \qquad j > k . \tag{4.73}$$

With this assumption, it will be found that the minimization of $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k \mathbf{e}_k})$ depends only on $\mathbf{B}_k$ and $\mathbf{F}_k$. Therefore, the solution of $\mathbf{B}$ and $\mathbf{F}$ that minimizes $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{ee}})$ can be found by minimizing $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k \mathbf{e}_k})$ with respect to $\mathbf{B}_k$ and $\mathbf{F}_k$ for $k = 1, \ldots, K$, separately. The matrix $\mathbf{F}_k$ that satisfies (4.73) can be represented as $\mathbf{F}_k = \mathbf{N}_k \mathbf{A}_k$. Here, $\mathbf{N}_k$ is an $N \times (N - \sum_{i=k+1}^{K} m_i)$ matrix whose columns are the orthonormal basis vectors of $\mathrm{null}(\bar{\mathbf{H}}_k)$, where

$$\bar{\mathbf{H}}_k \equiv [\mathbf{H}_{k+1}^T, \ldots, \mathbf{H}_K^T]^T . \tag{4.74}$$

It can be found by doing the SVD [61] on $\bar{\mathbf{H}}_k$. $\mathbf{A}_k$ can be any matrix of size $(N - \sum_{i=k+1}^{K} m_i) \times m_k$.

Using this assumption, it can be found that (4.72) is equivalent to

$$\mathbf{\Phi}_{\mathbf{e}_k \mathbf{e}_k} = \sigma_n^2 \bar{\mathbf{B}}_k (\bar{\mathbf{F}}_k^H \mathbf{H}_k^H \mathbf{H}_k \bar{\mathbf{F}}_k + \xi \mathbf{I})^{-1} \bar{\mathbf{B}}_k^H , \tag{4.75}$$

where

$$\bar{\mathbf{F}}_k \equiv [\mathbf{F}_k, \mathbf{F}_{k+1}, \ldots, \mathbf{F}_K] \tag{4.76}$$

and $\bar{\mathbf{B}}_k$ is made up of the $(\sum_{i=1}^{k-1} m_i + 1)$ th,...,Mth columns of $\mathbf{B}_k$, i.e., $\mathbf{B}_k \equiv [\mathbf{0}_{m_k, \sum_{i=1}^{k-1} m_i}, \bar{\mathbf{B}}_k]$. Now, $\bar{\mathbf{B}}_k$ that minimizes $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ can be found using the following equations:

$$\bar{\mathbf{F}}_k^H \mathbf{H}_k^H \mathbf{H}_k \bar{\mathbf{F}}_k + \xi \mathbf{I} = \bar{\mathbf{R}}_k^H \bar{\mathbf{R}}_k, \tag{4.77}$$

$$\bar{\mathbf{B}}_k = \mathbf{G}_k \bar{\mathbf{R}}_k', \tag{4.78}$$

$$\mathbf{G}_k = \mathrm{diag}[[\bar{\mathbf{R}}_k']_{11}^{-1}, [\bar{\mathbf{R}}_k']_{22}^{-1}, \ldots, [\bar{\mathbf{R}}_k']_{m_k m_k}^{-1}], \tag{4.79}$$

where (4.77) is the Cholesky factorization [61] of $\bar{\mathbf{F}}_k^H \mathbf{H}_k^H \mathbf{H}_k \bar{\mathbf{F}}_k + \xi \mathbf{I}$, and $\bar{\mathbf{R}}_k'$ is made up of the 1th,..., m_k th rows of $\bar{\mathbf{R}}_k$. The proof of $\bar{\mathbf{B}}_k$ found by (4.77)–(4.79) minimizing $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ is included in Appendix J. One can see that $\mathbf{G}_k$ in (4.77) –(4.79) can be also found by using the following equations

$$\mathbf{F}_k^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{F}_k + \xi \mathbf{I} = \mathbf{R}_k^H \mathbf{R}_k, \tag{4.80}$$

$$\mathbf{G}_k = \mathrm{diag}[[\mathbf{R}_k]_{11}^{-1}, [\mathbf{R}_k]_{22}^{-1}, \ldots, [\mathbf{R}_k]_{m_k m_k}^{-1}], \tag{4.81}$$

where (4.80) is the Cholesky factorization of $\mathbf{F}_k^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{F}_k + \xi \mathbf{I}$.

With the $\bar{\mathbf{B}}_k$ found from (4.77)–(4.79), $\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k}$ becomes

$$\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k} = \sigma_n^2 \mathbf{G}_k \mathbf{G}_k^H. \tag{4.82}$$

From (4.80)–(4.82), it can be seen that $\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k}$ is only influenced by $\mathbf{F}_k$, while $\mathbf{F}_j$ ($j \neq k$) have no influence on $\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k}$. Therefore, our final task is to design the pre-processing matrix $\mathbf{F}_k$ to minimize $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$. Since $\mathbf{F}_k$ is assumed to be $\mathbf{F}_k = \mathbf{N}_k \mathbf{A}_k$ and $\mathbf{N}_k$ is known, only $\mathbf{A}_k$ needs to be found. It is noted that $\mathbf{A}_k$ should make $\mathbf{F}_k$ satisfy the transmitted power constraint in (4.37). In this thesis, for simplicity only $\mathbf{A}_k$ that have orthonormal columns are considered, i.e., $\mathbf{A}_k$ should satisfy

$$\mathbf{A}_k^H \mathbf{A}_k = \mathbf{I}. \tag{4.83}$$

(4.83) directly implies that the transmitted power constraint (4.37) is satisfied. Simulation results in Section 4.3.3.2 will show that $\mathbf{A}_k$ found under this assumption can achieve very good performance.

$\mathbf{A}_k$ that satisfies (4.83) and minimizes $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$ is found to be

$$\mathbf{A}_k = \mathbf{V}_k \mathbf{P}_k \tag{4.84}$$

where $\mathbf{V}_k$ is found from the SVD of $\mathbf{H}_k\mathbf{N}_k$, i.e.,

$$\mathbf{H}_k\mathbf{N}_k = \mathbf{U}_k \begin{bmatrix} \mathbf{\Sigma}_k & \mathbf{0}_{m_k, N-\sum_{i=k}^{K} m_i} \end{bmatrix} \begin{bmatrix} \mathbf{V}_k^H \\ \tilde{\mathbf{V}}_k^H \end{bmatrix}, \tag{4.85}$$

and $\mathbf{P}_k$ is found by the GMD [90], [94] (or the equal-diagonal QR decomposition [89]) of $\mathbf{J}_k$, i.e.,

$$\mathbf{J}_k = \mathbf{Q}_k \mathbf{R}_k^J \mathbf{P}_k^H \tag{4.86}$$

where

$$\mathbf{J}_k \equiv \begin{bmatrix} \mathbf{\Sigma}_k \\ \sqrt{\xi}\mathbf{I} \end{bmatrix}. \tag{4.87}$$

The GMD of a matrix has been described in Section 4.3.2. Using $\mathbf{A}_k$ found by (4.84)–(4.87), the variances of $e_k^{(1)}, e_k^{(2)}, \ldots, e_k^{(m_k)}$, which are the errors of MS_k's received signals, are

$$\sigma_{e_k^{(1)}}^2 = \sigma_{e_k^{(2)}}^2 = \ldots = \sigma_{e_k^{(m_k)}}^2 = \sigma_n^2 \left(\prod_{i=1}^{m_k} \sigma_k^{(i)2}\right)^{-1/m_k} \tag{4.88}$$

where $\sigma_k^{(i)}$ $(i = 1, \ldots, m_k)$ are the singular values of $\mathbf{J}_k$. Since $\sigma_{e_k^{(1)}}^2, \sigma_{e_k^{(2)}}^2, \ldots, \sigma_{e_k^{(m_k)}}^2$ are all equal, we use one symbol $\sigma_{e_k}^2$ to represent this value. Proof that $\mathbf{A}_k$ given by (4.84)–(4.87) minimizes $\mathrm{tr}(\mathbf{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$, and proof of (4.88) are included in Appendix K.

It is noted that when all $\mathbf{A}_k$ $(k = 1, \ldots, K)$ are found through (4.84)–(4.87), the matrix $\mathbf{F}$ is defined. Then, matrices $\mathbf{B}$ and $\mathbf{D}$ can be derived from (4.77)–(4.79) and (4.71).

For the above solution found for $\mathbf{A}_k$, it is interesting to find that the pre-processing matrix $\mathbf{F}$ in the MMSE nonlinear joint Tx-Rx processing algorithm satisfies $\mathbf{F}^H\mathbf{F} = \mathbf{I}$ (see Appendix L for the proof), which means $\mathbf{F}$ has orthonormal columns. Therefore, one can see that the transmitted power is allocated equally to all the data streams. As we have discussed in Section 2.5.1, due to this feature, when applied to the sytems with $M = N$, the pre-processing in the MMSE nonlinear joint Tx-Rx processing algorithm does not influence the required dynamic range of the transmitted power at each antenna. The transmitted powers at the multiple transmit antennas will all have the same value and this value will not change as the channel changes. This feature is a clear advantage in the design of a practical system.

It is noted that when the MMSE nonlinear pre-processing algorithm proposed in Section 2.5 is used, the transmitted power is also allocated equally to all the data streams. Comparing the MMSE nonlinear joint Tx-Rx processing algorithm and the MMSE nonlinear pre-processing algorithm, we find that by effectively taking into consideration the processing ability of the mobiles, the MMSE nonlinear joint Tx-Rx processing algorithm we have introduced above can achieve significantly better performance than the MMSE nonlinear pre-processing algorithm. The performance comparison of these two algorithms is included in Section 4.3.3.2. In addition, it can be found that although the computational complexity of the MMSE nonlinear joint Tx-Rx processing algorithm is higher than that of the MMSE nonlinear pre-processing algorithm, the computational complexity of the former is still of the same order as that of the latter.

The design of the MMSE-criterion-based nonlinear joint Tx-Rx processing algorithms using the same structure as shown in Figure 4.7 has also been studied in some other recent work [80], [97]. The different design and advantage of the MMSE nonlinear joint Tx-Rx processing algorithm proposed in this thesis can be easily seen. The algorithm of [97] is developed based on the duality of multi-user MIMO downlink with dirty-paper coding and multi-user MIMO uplink with perfect SIC. It is designed to minimize the total transmitted power while satisfying individual MSE (or SINR) target at each mobile. A different design

goal of minimizing the sum MSE at all the receive branches under a fixed total transmitted power is achieved by the proposed MMSE nonlinear joint Tx-Rx processing algorithm and the work in [80]. In the algorithm of [80] only the MUI is pre-eliminated by nonlinear pre-processing, while the self-interference of each user is processed by linear joint Tx-Rx processing. The proposed MMSE nonlinear joint Tx-Rx processing algorithm can be seen as using nonlinear pre-processing to pre-eliminate the MUI, while the self-interference is suppressed via joint nonlinear pre-processing and linear receiver processing, which makes the design of the proposed algorithm different from that of [80]. Moreover, the processing matrices in [80] and [97] need to be found by iterative methods, while closed-form expressions have been found for the processing matrices in the proposed algorithm. This makes the proposed MMSE nonlinear joint Tx-Rx processing algorithm more attractive for practical implementation.

4.3.3.1 Ordering in the MMSE Nonlinear Joint Tx-Rx Processing Algorithm

When the MMSE nonlinear joint Tx-Rx processing algorithm is used, it can be found that the channel matrices of different mobiles, i.e., $\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_K$ in $\mathbf{H}$, can be ordered to achieve better system performance. The change of the system structure due to ordering should be the same as that for the ZF nonlinear joint Tx-Rx processing algorithm described in Section 4.3.2.3.1. We still use $L \equiv \{l_1, l_2, ..., l_K\}$ to represent an order of these channel matrices. If the order of $\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_K$ is found to be $L \equiv \{l_1, l_2, ..., l_K\}$, $\mathbf{H}$ should be ordered into $\mathbf{H}^{(L)} = [\mathbf{H}_{l_1}^T, \mathbf{H}_{l_2}^T, ..., \mathbf{H}_{l_K}^T]^T$. Then, processing matrices $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{D}$ should be generated using $\mathbf{H}^{(L)}$ instead of $\mathbf{H}$, and they are denoted as $\mathbf{B}^{(L)}$, $\mathbf{F}^{(L)}$, and $\mathbf{D}^{(L)}$. Next, let us study how to design the ordering method for the MMSE nonlinear joint Tx-Rx processing algorithm to improve performance.

For this algorithm, if order $L \equiv \{l_1, l_2, ..., l_K\}$ is used, since $\mathbf{B}^{(L)}$ is assumed to be a unit upper triangular matrix, the data vectors will be precoded from $\mathbf{a}_{l_K}, \mathbf{a}_{l_{K-1}}$, to $\mathbf{a}_{l_1}$. It can be found that the performance of MS_{l_j} is better than the performance of MS_{l_k} if $j > k$. In other words, the mobile whose data is precoded early has

better performance than the mobile whose data is precoded late. Since the performance of the worse mobile has greater influence on the overall system performance, the mobile that is precoded late affects the overall system performance more significantly. Therefore, the "best-first" ordering described in Section 3.3, which always tries to order the mobile with the best performance (smallest error variance) to the latest precoding position, is a reasonable ordering method for the MMSE nonlinear joint Tx-Rx processing algorithm. The detailed explanation of the working mechanism of the "best-first" ordering can be found in Section 3.3. In the following a brief description of the "best-first" ordering for the MMSE nonlinear joint Tx-Rx processing algorithm is given.

During the "best-first" ordering, the final order $L \equiv \{l_1, l_2, ..., l_K\}$ is found from l_1, l_2, to l_K iteratively. To find l_i, all the unordered mobiles (the mobiles whose indices have not been chosen as $l_1,...,l_{i-1}$) are tested. The index of the mobile that will lead to the minimum error variance $\sigma^2_{e_{l_i}}$ will be ordered into the position l_i.

Simulation results in next section will show that the "best-first" ordering method can greatly improve the performance of the MMSE nonlinear joint Tx-Rx processing algorithm.

4.3.3.2 Simulation Results

To show the advantage of the MMSE nonlinear joint Tx-Rx processing algorithm proposed in this section, we compare it with other pre-processing or joint Tx-Rx algorithms. The algorithms we compare include: i) the MMSE nonlinear pre-processing algorithm proposed in Section 2.5. ii) MLD [30]–[32], MMSE DFD with the "best-first" ordering [28], or MMSE linear detector (LD) [30], combined with the L-DECOM technique described in Section 4.2.1, whose corresponding curves are denoted as L-DECOM MLD, L-DECOM MMSE DFD, and L-DECOM MMSE LD. The L-DECOM MMSE LD is a linear joint Tx-Rx processing algorithm for the downlink of multi-user MIMO systems, while the L-DECOM MMSE DFD and L-DECOM MLD are nonlinear joint Tx-Rx processing algorithms for the downlink of multi-user MIMO systems, which are made up of

linear pre-processing and nonlinear receiver processing. iii) The ZF nonlinear joint Tx-Rx processing algorithm proposed in Section 4.3.2. It can be found that all these algorithms (including the MMSE nonlinear pre-processing algorithm, L-DECOM MLD, L-DECOM MMSE DFD, L-DECOM MMSE LD, and the ZF nonlinear joint Tx-Rx processing algorithm) realize equal transmitted power allocation to all the data streams.

In our simulation, BER vs. SNR curves are determined. QPSK with Gray encoding [74] is used. The BER is the average BER of all the mobiles. The SNR is defined by (2.43). The flat Rayleigh fading channel is assumed to change independently from one symbol interval to the next.

Figure 4.13 presents performance comparison in the downlink of a multi-user MIMO system with $N = 6$, $K = 2$, and $m_1 = m_2 = 3$. The "best-first" ordering method is applied to the MMSE nonlinear joint Tx-Rx processing algorithm, the ZF nonlinear joint Tx-Rx processing algorithm, and the MMSE nonlinear pre-processing algorithm. It is clearly seen that the MMSE nonlinear joint Tx-Rx processing algorithm can achieve significantly better performance than the other algorithms. One should note that the MMSE nonlinear joint Tx-Rx processing algorithm even outperforms the L-DECOM MLD algorithm, which has much higher computational complexity. At BER = 10^{-3} the MMSE nonlinear joint Tx-Rx processing algorithm achieves approximately 15 dB gain over the L-DECOM MMSE LD algorithm, 5 dB gain over the L-DECOM MMSE DFD algorithm, 3.4 dB gain over the MMSE nonlinear pre-processing algorithm, 3 dB gain over the ZF nonlinear joint Tx-Rx processing algorithm, and 2.2 dB gain over the L-DECOM MLD algorithm. The benefits of the MMSE nonlinear joint Tx-Rx processing algorithm over the other algorithms can be explained as follows. The advantage of the MMSE nonlinear joint Tx-Rx processing algorithm over the MMSE nonlinear pre-processing algorithm lies in that it effectively utilizes the processing capabilities of the mobiles. Compared to the MMSE nonlinear joint Tx-Rx processing algorithm, the performance of the L-DECOM based algorithms is limited by the linear pre-processing applied. The MMSE nonlinear joint Tx-Rx processing algorithm can outperform the ZF nonlinear joint Tx-Rx processing

algorithm because the application of the MMSE criterion reduces the noise enhancement.

Performance comparison of different algorithms in the downlink of a multi-user MIMO system with $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$ is given in Figure 4.14. The "best-first" ordering method is applied to the MMSE nonlinear joint Tx-Rx processing algorithm, the ZF nonlinear joint Tx-Rx processing algorithm, and the MMSE nonlinear pre-processing algorithm. The advantage of the MMSE nonlinear joint Tx-Rx processing algorithm can also be seen clearly. At $\mathrm{BER} = 10^{-3}$ the MMSE nonlinear joint Tx-Rx processing algorithm achieves approximately 15 dB gain over the L-DECOM MMSE LD algorithm, 9 dB gain over the L-DECOM MMSE DFD algorithm, 5 dB gain over the L-DECOM MLD algorithm, 3.6 dB gain over the ZF nonlinear joint Tx-Rx processing algorithm, and 1.5 dB gain over the MMSE nonlinear pre-processing algorithm.

Figure 4.15 shows performance comparison of the MMSE nonlinear joint Tx-Rx processing algorithm without ordering or with the "best-first" ordering described in Section 4.3.3.1 in the downlink of a multi-user MIMO system with $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$. One can see that the "best-first" ordering significantly improves the system performance. At BER = 10^{-3}, using the "best-first" ordering, approximately 2 dB gain is obtained for the MMSE nonlinear joint Tx-Rx processing algorithm.

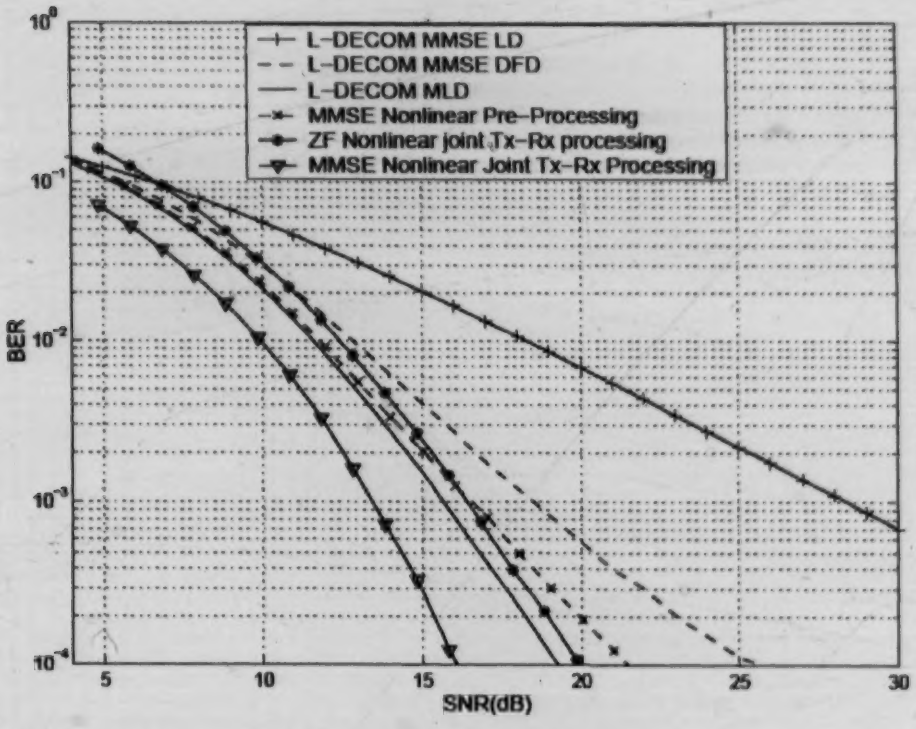


Figure 4.13 Performance of the MMSE nonlinear joint Tx-Rx processing algorithm (QPSK, $N = 6$, $K = 2$, and $m_1 = m_2 = 3$).

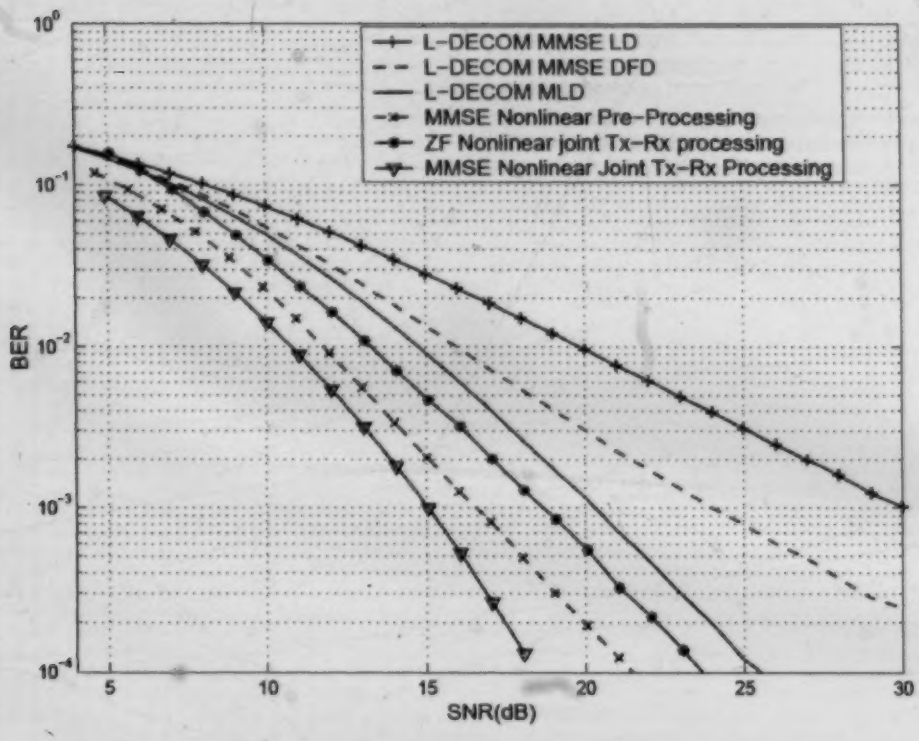


Figure 4.14 Performance of the MMSE nonlinear joint Tx-Rx processing algorithm (QPSK, $N = 6$, $K = 3$, and $m_1 = m_2 = m_3 = 2$).

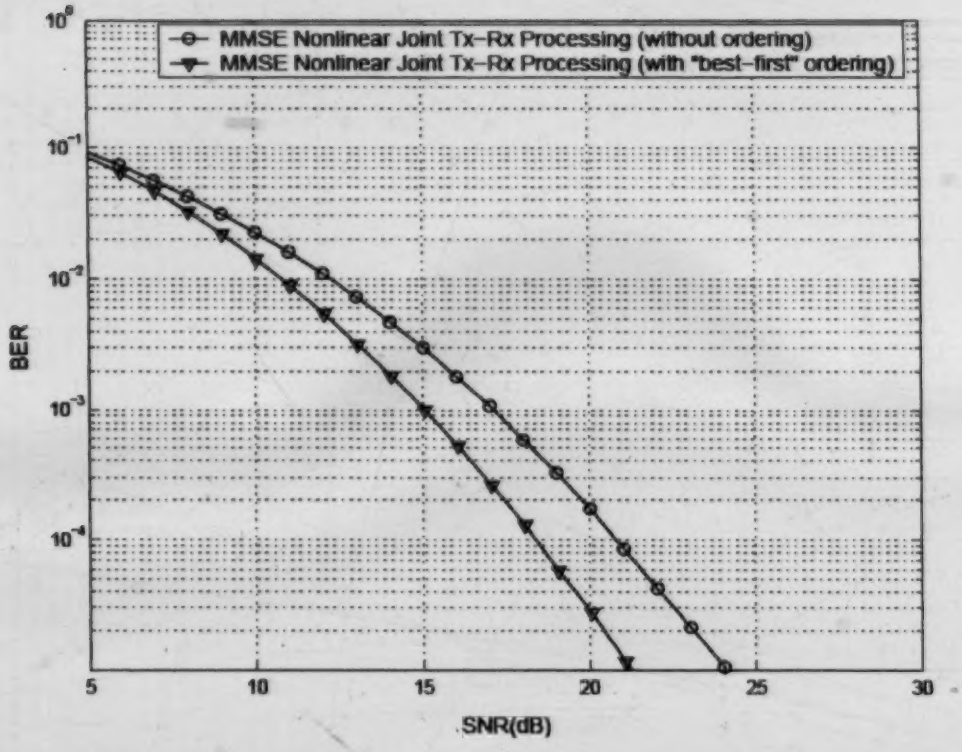


Figure 4.15 Performance comparison of the MMSE nonlinear joint Tx-Rx processing algorithm when no ordering or the "best-first" ordering is used (QPSK, $N = 6, K = 3$, and $m_1 = m_2 = m_3 = 2$).

Chapter 5 Conclusions

5.1 Conclusions

This thesis is focused on the design of spatial layer separation algorithms for the downlink of multi-user MIMO systems. Due to its abilities of achieving high sum capacity and accommodating multiple users, the multi-user MIMO systems have attracted a lot research interests recently. One major problem in the multi-user MIMO systems is the spatial layer separation for the downlink since the mobiles are decentralized and uncoordinated. In this thesis, several nonlinear pre-processing and nonlinear joint Tx-Rx processing algorithms have been proposed to solve this problem, and the proposed algorithms have shown significantly better performance than the previous work in this field. The major contributions of this thesis are summarized in this section.

In the first chapter, a brief introduction to the MIMO antenna systems and the multi-user MIMO systems has been given as background knowledge. The work in the literature on the pre-processing and joint Tx-Rx processing algorithms for the multi-user MIMO downlink has been summarized. Since the THP can effectively constrain the transmitted power increase, nonlinear pre-processing applying the THP can achieve significantly better performance than linear pre-processing. Therefore, our research has focused on the design of the nonlinear pre-processing and nonlinear joint Tx-Rx processing algorithms for the multi-user MIMO downlink. The organization of this thesis has also been introduced in this chapter.

Chapter 2 is focused on the design of the nonlinear pre-processing algorithms for the multi-user MIMO downlink. The TH precoder, which is an important part of the nonlinear pre-processing algorithms, is first introduced. Then, we propose a generalized ZF nonlinear pre-processing structure which can realize ZF-based

nonlinear pre-processing algorithm that achieves minimum BER for each mobile under any given power allocation to different data streams. Based on this structure, we have designed the ZF nonlinear pre-processing algorithm, which achieves minimum BER at each mobile under any given relative SNR requirement, the ZF nonlinear pre-processing algorithm, which achieves minimum total transmitted power while satisfying the individual SNR target at each mobile, and the minimum BER ZF nonlinear pre-processing algorithm, which achieves minimum average BER of the system. Since the MMSE criterion can lead to better performance than the ZF criterion due to the avoidance of the noise enhancement, we also designed the MMSE criterion based nonlinear pre-processing algorithm for the multi-user MIMO downlink in this chapter. Simulation results showed that the MMSE nonlinear pre-processing algorithm can achieve significantly better performance than the ZF criterion based counterpart. It should be noted that the nonlinear pre-processing algorithms proposed in this chapter can also be used in other types of multiple input multiple output systems that are made up of a centralized transmitter and several decentralized receivers, such as the downlink of the CDMA systems and the downstream of the "MIMO-DSL" systems[4] [98], [99].

Since the performance of the nonlinear pre-processing algorithms employing the THP is influenced by ordering the rows of the channel matrix, there is a problem of finding the suitable order to improve the performance of these algorithms. The ordering problems of nonlinear pre-processing algorithms were studied in Chapter 3. The most important contribution of this chapter is the introduction of the optimal ordering lemma, which gives the conditions under which the "best-first" ordering method is optimal. This lemma has then been used to solve the ordering problems of the conventional ZF nonlinear pre-processing algorithm [51], [52], the MMSE nonlinear pre-processing algorithm proposed in Chapter 2, and the ZF nonlinear joint Tx-Rx processing algorithm proposed in

[4] In MIMO-DSL systems, transmission lines of multiple DSL customers are exclusively controlled by a single entity. Thus, pre-processing algorithms or signal detection algorithms can be applied to cancel the crosstalk in the downstream or upstream. The detailed introduction of the MIMO-DSL systems can be found in [98].

Chapter 4. The ordering issues for other ZF-based nonlinear pre-processing algorithms proposed in Chapter 2 are also studied in this chapter, and it is found that the "best-first" ordering method can achieve very satisfactory performance for these algorithms.

The spatial layer separation algorithms for the downlink of multi-user MIMO systems with multiple-antenna mobiles have been considered in Chapter 4. The nonlinear joint Tx-Rx processing algorithms we propose utilize the benefit of the nonlinear pre-processing employing THP to constrain the transmitted power increase and the additional degree of freedom provided by the deployment of multiple antennas at the mobiles, so better performance than that of the linear joint Tx-Rx processing algorithms and the pre-processing only algorithms can be achieved. The first major contribution of this chapter is the introduction of the NL-DECOM technique. This technique applies nonlinear pre-processing to decompose a multi-user MIMO channel into multiple independent single-user MIMO channels, so the data streams of each mobile can be seen as passing through the decomposed equivalent single-user MIMO channel without any MUI. Operating jointly with different single-user MIMO spatial layer separation algorithms, this technique realizes different nonlinear joint Tx-Rx processing algorithms for the multi-user MIMO downlink with multiple-antenna mobiles. It can achieve significantly better BER performance, higher maximum achievable sum rate, and higher system design flexibility than the L-DECOM technique [57]–[60] known in the literature. In the NL-DECOM technique, the nonlinear pre-processing is used to pre-eliminate only the MUI. To achieve better performance, the ZF and MMSE nonlinear joint Tx-Rx processing algorithms that use nonlinear pre-processing to pre-eliminate both the MUI and the self-interference at the transmitter are proposed. Linear receiver-processing is applied in these algorithms to utilize the processing capabilities of the mobiles without increasing the complexity of the mobiles too much. The performance advantage of these two algorithms has been shown by simulations. It has been found that if the channel matrices of different mobiles are ordered suitably, better performance can be achieved for the nonlinear joint Tx-Rx processing algorithms proposed in

this chapter. The ordering problems of these algorithms have also been studied and solved in this chapter.

5.2 Suggestions for Future Work

All the nonlinear pre-processing algorithms and the nonlinear joint Tx-Rx processing algorithms proposed in this thesis are designed under the assumption that perfect CSI is known at the transmitter. The CSIT can be obtained by channel estimation in reciprocal TDD systems or by feedback from the receivers. However, in practical systems, sometimes it is difficult for the transmitter to know the complete CSIT. For the TDD transmission, the CSIT estimated from the previous frame may be not precise for the current frame, if the radio channel changes sufficiently fast. If feedback channel is used, the feedback of the CSI may lead to CSI errors or make the CSI outdated. In addition, the limitation on the bandwidth of the feedback channel also constrains the precision of the CSIT. Hence, one direction of the future work can be the design of nonlinear pre-processing algorithms and nonlinear joint Tx-Rx processing algorithms using partial CSIT. There are two types of models for the partial CSIT, which have been generally used in the literature and are of interest. The first type of model assumes that the second order statistics of the channel (long term CSI) are available at the transmitter [100]–[102]. The second type of model assumes that the partial CSIT is the quantized value of the real CSIT [103], [104]. This model covers the scenario that the partial CSIT is provided by feedback channel with limited bandwidth. If this model is used, beside the design of the layer separation algorithms, the design of the vector quantization that could cooperate well with the layer separation algorithms is also an important problem.

The nonlinear pre-processing algorithms proposed in Chapter 2 and the ZF/MMSE nonlinear joint Tx-Rx processing algorithm proposed in Chapter 4 are designed under the assumption that the number of antennas at the base station is no less than the total number of antennas at the mobiles simultaneously communicating with the base station. The NL-DECOM technique proposed in

Section 4.2 requires that $N > \sum_{i=2}^{K} m_i$, where N is the number of antennas at the base station, m_i is the number of antennas at mobile i, and there are K mobiles simultaneously communicating with the base station. In systems where there are a large number of mobiles, the above assumption can be satisfied by mobile subset selection (scheduling) schemes, which select a subset of mobiles for communication with the base station simultaneously. The mobile subset selection schemes are another direction of possible future research work. If designed properly, mobile scheduling schemes are very beneficial for the system performance since the multi-user diversity in wireless communication systems can be utilized. For example, the maximum achievable sum rate of the multi-user MIMO downlink when the conventional ZF nonlinear pre-processing algorithm with mobile scheduling is applied has been analyzed in [105]. It has been shown there that if the multi-user diversity is utilized properly the maximum achievable sum rate of the multi-user MIMO downlink using the conventional ZF nonlinear pre-processing algorithm is higher than the theoretical capacity of the corresponding point-to-point MIMO antenna system and the maximum achievable sum rate of the V-BLAST system using the ZF DFD, as long as the number of mobiles in the system is sufficiently large [105]. However, it should be noted that the scheduling method of [105] is very complex. Design of the mobile scheduling schemes for the nonlinear pre-processing algorithms proposed in Chapter 2 and the nonlinear joint Tx-Rx processing algorithms proposed in Chapter 4 may be of interest in the future. Fairness as well as the performance/complexity tradeoff are important factors that should be considered in the design of a practical mobile scheduling scheme.

The ZF and MMSE nonlinear joint Tx-Rx processing algorithms proposed for the downlink of multi-user MIMO systems with multiple-antenna mobiles in Chapter 4 have been developed based on the assumption that the number of data streams transmitted to each scheduled mobile is equal to the number of antennas at that mobile. Thus, the number of scheduled mobiles, K, simultaneously

communicating with the base station is restricted by $N \geq \sum_{i=1}^{K} m_i$. Also, sometimes it may be desirable that the number of data streams transmitted to each scheduled mobile is less than the number of antennas at that mobile. Such a configuration can allow a larger number of mobiles simultaneously communicating with the base station. For example, at most N mobiles can be accommodated simultaneously by the base station with N antennas. Linear joint Tx-Rx processing algorithms under the assumption that the number of data streams transmitted to each mobile can be less than the number of antennas have been proposed in the literature, e.g. in [53]–[56]. The design of the nonlinear joint Tx-Rx processing algorithms under the same assumption is still an open question and is one possible direction of future work.

References

[1] "Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000," *RECOMMENDATION ITU-R M.1645*, June 2003

[2] Y. Kim, B. J. Jeong, J. Chung, C. Hwang, J. S. Ryu, K. Kim, and Y. K. Kim, "Beyond 3G: Vision, requirements, and enabling technologies," *IEEE Commun. Mag.*, vol. 41, pp. 120–124, Mar. 2003.

[3] S. Ryu, D. Oh, G. Sihn, and K. Han, "Perspective of the next generation mobile communications and services," in *Proc. PIMRC'04*, Sep. 2004, pp. 643–647.

[4] E. D. Re and L. Pierucci, "Multiple antenna systems: Frontier for wireless access," in *Proc. PIMRC'04*, Sep. 2004, pp. 930–934.

[5] D. G. Brennan, "Linear diversity combining techniques," in *Proc. IRE*, vol. 47, pp. 1075–1102, June 1995 (Reprinted in *Proceedings of the IEEE*, vol. 91, pp. 331–356, Feb. 2003).

[6] G. L. Stüber, *Principles of Mobile Communication.* 2nd ed., Boston: Kluwer Academic Publishers, 2001.

[7] M. Chryssomallis, "Smart antennas," *IEEE Antennas Propagat. Mag.*, vol. 42, pp. 129–136, June 2000.

[8] J. H. Winters, "Smart antennas for wireless systems," *IEEE Pers. Commun.*, vol. 5, pp. 23–27, Feb. 1998.

[9] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, Mar. 1998.

[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Communications*, vol. 10, no. 6, pp. 585–595, 1999.

[11] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multielement antennas," *Bell Labs Tech. J.*, pp. 41–59, Autumn 1996.

[12] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.

[13] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 281–302, Apr. 2003.

[14] H. Yao and G. W. Wornell, "Structured space-time block codes with optimal diversity-multiplexing tradeoff and minimum delay," in *Proc. GLOBECOM'03*, vol. 4, pp. 1941–1945.

[15] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1121–1128, May 1999.

[16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge: Cambridge University Press, 2005.

[17] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *IEE Elect. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[18] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE'98*, pp. 295–300.

[19] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[20] S. Alamouti, "A simple transmitter diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[21] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[22] W. Rhee and J. M. Cioffi, "On the capacity of multiuser wireless channels with multiple antennas," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2580–2595, Oct. 2003.

[23] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684–702, June 2003.

[24] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1691–1706, July 2003.

[25] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2658–2668, Oct. 2003.

[26] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[27] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.

[28] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. ICASSP'00*, Istanbul, Turkey, June 2000, pp. II737–II740.

[29] G. Ginis and J. M. Cioffi, "On the relation between V-BLAST and the GDFE," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 364–366, Sep. 2001.

[30] B. A. Bjerke and J. G. Proakis, "Multiple-antenna diversity techniques for transmission over fading channels," in *Proc. WCNC'99*, vol. 3, pp. 1038–1042.

[31] X. Zhu and R. D. Murch, "Performance analysis of maximum likelihood detection in a MIMO antenna system," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 187–191, Feb. 2002.

[32] R. van Nee, A. van Zelst, and G. Awater, "Maximum likelihood decoding in a space division multiplexing system," in *Proc. VTC'00-Spring*, vol. 1, pp. 6–10.

[33] A. Bhargave, R. J. P. de Figueiredo, and T. A Eltoft, "A detection algorithm for the V-BLAST system," in *Proc. GLOBECOM'01*, vol. 1, pp. 494–498.

[34] W. J. Choi, R. Negi, and J. M. Cioffi, "Combined ML and DFE decoding for the V-BLAST system," in *Proc. ICC'00*, vol. 3, pp. 1243–1248.

[35] Y. Huang, J. Zhang, and P. M. Djuric, "Detection with particle filtering in BLAST systems," in *Proc. ICC'03*, vol. 4, pp. 2306–2310.

[36] M. O. Damen, K. Abed-Meraim, and S. Burykh, "Iterative QR detection for BLAST," *Wireless Personal Communications*, vol. 19, No. 3, pp. 179–192, Dec. 2001.

[37] X. Li, H. C. Huang, A. Lozano, and G. J. Foschini, "Reduced-complexity detection algorithms for systems using multi-element arrays," in *Proc. GLOBECOM'00*, vol. 2, pp. 1072–1076.

[38] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Trans. Signal Processing*, vol. 51, pp. 1722–1730, July 2003.

[39] A. Sezgin, E. A. Jorswieck, and V. Jungnickel, "Maximum diversity detection for layered space-time codes," in *Proc. VTC'03-Spring*, vol. 2, pp. 833–837.

[40] B. R. Vojcic and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Commun.*, vol. 46, pp. 1346–1355, Oct. 1998.

[41] A. N. Barreto and G. Fettweis, "Capacity increase in the downlink of spread spectrum systems through joint signal precoding," in *Proc. ICC'01*, vol. 4, pp. 1142–1146.

[42] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter for the downlink of TDD DS-CDMA systems," in *Proc. ISSSTA'02*, vol. 1, pp. 9–13.

[43] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Processing*, vol. 53, pp. 2700–2712, Aug. 2005.

[44] J. Liu and W. A. Krzymień, "Signal detection and pre-processing algorithms in layered space-time multiple input multiple output antenna systems," in *Proc. Wireless'04*, Calgary, Canada, July 2004, pp. 9–21.

[45] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[46] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice-strategies for cancelling known interference," in *Proc. ISITA'00*, Honolulu, HI, Nov. 2000.

[47] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice-strategies for cancelling known interference," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3820–3833, Nov. 2000.

[48] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electron. Lett.*, vol. 7, pp. 138–139, Mar. 1971.

[49] H. Harashima and H. Miyakawa, "A method of code conversion for digital communication channels with intersymbol interference," *Transactions of the Institute of Electronics and Communications Engineers of Japan*, pp. 272–273, Jun. 1969 (in Japanese).

[50] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, pp. 774–780, Aug. 1972.

[51] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. New York: John Wiley & Sons, 2002.

[52] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1305–1316, July 2004.

[53] Z. Pan, K. K. Wong, and T. S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1969–1973, Nov. 2004.

[54] K. K. Wong, R. D. Murch, and K. B. Letaief, "A joint-channel diagonalization for multiuser MIMO antenna systems," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 773–786, July 2003.

[55] A. J. Tenenbaum and R. S. Adve, "Joint multiuser transmit-receive optimization using linear processing," in *Proc. ICC'04*, Paris, France, June 2004, pp. 588–592.

[56] J. Zhang, Y. Wu, S. Zhou, and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Commun. Lett.*, vol. 9, pp. 991–993, Nov. 2005.

[57] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 461–471, Feb. 2004.

[58] M. Rim, "Multi-user downlink beamforming with multiple transmit and receive antennas," *Electron. Lett.*, vol. 38, pp. 1725–1726, Dec. 2002.

[59] L. U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 20–24, Jan. 2004.

[60] A. Bourdoux and N. Khaled, "Joint TX-RX optimisation for MIMO-SDMA based on a null-space constraint," in Proc. *VTC'02-Fall*, Vancouver, Canada, Sep. 2002, pp. 171–174.

[61] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, 1996.

[62] T. Haustein, M. Schubert, and H. Boche, "On power reduction strategies for the multi-user downlink with decentralized receivers," in *Proc. VTC'03-Spring*, Jeju, Korea, Apr. 2003, pp. 1007–1011.

[63] S. Verdu, *Multiuser Detection*. Cambridge: Cambridge University Press, 1998.

[64] A. Klein, G. K. Kaleh, and P. W. Baier, "Zero forcing and minimum mean-square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 276–287, May 1996.

[65] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Trans. Signal Processing*, vol. 53, pp. 2324–2334, July 2005.

[66] M. Schubert and S. Shi, "MMSE transmit optimization with interference pre-compensation," in *Proc. VTC'05-Spring*, Stockholm, Sweden, pp. 845–849, May 2005.

[67] "IEEE Standard for local and metropolitan area networks; Part 16: Air interface for fixed and mobile broadband wireless access systems," *IEEE 802.16e Standard*, http://standards.ieee.org/getieee802/802.16.html, Feb. 2006.

[68] A.Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, and J. McGeehan, "A comparison of the HIPERLAN/2 and IEEE 802.11a wireless LAN standards," *IEEE Commun. Mag.*, pp. 172–180, May 2002.

[69] "3rd generation partnership project; Technical specification group radio access network; Base station (BS) radio transmission and reception (TDD)," 3GPP Technical Specification 25.105, V7.3.0, ftp://ftp.3gpp.org/specs/2006-09/Rel-7/25_series/, Oct. 2006.

[70] J. E. Mazo and J. Salz, "On the transmitted power in generalized partial response," *IEEE Trans. Commun.*, vol. 24, pp. 348–351, Mar. 1976.

[71] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. dissertation, MIT, Cambridge, MA, May 1989.

[72] M. S. Bartlett, "On the theory of statistical regression," in *Proc. R. Soc. Edinb. 53*, 1993, pp. 260–283.

[73] A. Duel-Hallen, "Decision-feedback multiuser detector for synchronous code-division multiple access channel," *IEEE Trans. Commun.*, vol. 41, no. 2, pp. 285–290, Feb. 1993.

[74] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2001.

[75] Z. Yan, K. M. Wong, and Z. Luo, "Optimal diagonal precoder for multiantenna communication systems," *IEEE Trans. Signal Processing*, vol. 53, pp. 2089–2100, June 2005.

[76] M. Payaro, A. Perez-Neira, and M. A. Lagunas, "Achievable rates for generalized spatial Tomlinson-Harashima precoding in MIMO systems," in *Proc. VTC'04-Fall*, Los Angeles, USA, Sep. 2004, pp. 2462–2466.

[77] M. Joham, J. Brehmer, and W. Utschick, "MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding," in *Proc. ITG SCC'04*, Erlangen, Germany, Jan. 2004, pp. 387–394.

[78] M. Joham, J. Brehmer, A. Voulgarelis, and W. Utschick, "Multiuser spatio-temporal Tomlinson-Harashima precoding for frequency selective vector channels," in *Proc. ITG Workshop on Smart Antennas*, Munich, Germany, Mar. 2004.

[79] C. Windpassinger, *Detection and Precoding for Multiple Input Multiple Output Channels*, Ph.D. thesis, Shaker Verlag, Aachen, Germany, 2004.

[80] A. Mezghani, R. Hunger, M. Joham, and W. Utschick, "Iterative THP transceiver optimization for multi-user MIMO systems based on weighted sum-MSE minimization," in *Proc. SPAWC'06*, Cannes, France, July 2006.

[81] Y. Jiang, X Zheng, and J. Li, "Asymptotic performance analysis of V-BLAST," in *Proc. GLOBECOM'05*, vol. 6, pp. 3882–3886.

[82] Y. Jiang, M. K. Varanasi, X Zheng, and J. Li, "Performance analysis of V-BLAST at high SNR regime," http://dsp.colorado.edu/~yjiang/papers/VBitwc.pdf.

[83] J. Liu and W. A. Krzymień, "A novel nonlinear precoding algorithm for the downlink of multiple antenna multi-user systems," in *Wireless Personal Communications*, vol. 2, Apr. 2007, pp. 207–223.

[84] R. D. Wesel and J. M. Cioffi, "Achievable rates for Tomlinson-Harashima precoding," *IEEE Trans. on Information Theory*, vol. 44, no. 2, pp. 824–831, Mar. 1998.

[85] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.

[86] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 773–775, Feb.-Mar.-Apr. 1995.

[87] M. A. Khalighi, J. M. Brossier, G. V. Jourdain, and K. Raoof, "Water filling capacity of Rayleigh MIMO channels," in *Proc. PIMRC'01*, San Diego, USA, Sep. 2001, pp. A-155 –A-158.

[88] X. Zhang and B. Ottersten, "Power allocation and bit loading for spatial multiplexing in MIMO systems," in *Proc. ICASSP'03*, Hong Kong, China, Apr. 2003, pp. V-53–V-56.

[89] J. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inform. Theory*, vol. 51, pp. 154–172, Jan. 2005.

[90] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Processing*, vol. 53, pp. 3791–3803, Oct. 2005.

[91] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[92] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.

[93] J. Liu and W. A. Krzymień, "Improved Tomlinson-Harashima precoding for the downlink of multiple antenna multi-user systems," in *Proc. WCNC'05*, New Orleans, USA, Mar. 2005, pp. 466–472.

[94] Y. Jiang, W. W. Hager, and J. Li, "The geometric mean decomposition," *Linear Algebra and Its Applications*, pp. 373–384, Feb. 2005.

[95] Y. Jiang, J. Li, and W. W. Hager, "Uniform channel decomposition for MIMO communications," *IEEE Trans. Signal Processing*, vol. 53, pp. 4283–4294, Nov. 2005.

[96] S. Loyka and F. Gagnon, "Performance analysis of the V-BLAST algorithm: An analytical approach," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1326–1337, July 2004.

[97] R. Doostnejad, T. J. Lim, and E. Sousa, "Precoding for the MIMO broadcast channels with multiple antennas at each receiver," in *Proc. of The Johns Hopkins University Conf. on Information Sciences and Systems*, Baltimore, MD, Mar. 2005.

[98] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1085–1104, June 2002.

[99] R. Cendrillon, M. Moonen, R. Suciu, and G. Ginis, "Simplified power allocation and TX/RX structure for MIMO-DSL," in *Proc. GLOBECOM'03*, vol. 4, pp. 1842–1846.

[100] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Linear and nonlinear preequalization/equalization for MIMO systems with long-term channel state information at the transmitter," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 373–378, Mar. 2004.

[101] M. T. Ivrlac, R. L. U. Choi, R. D. Murch, and J. A. Nossek, "Effective use of long-term transmit channel state information in multi-user MIMO communication systems," in *Proc. VTC'03-Fall*, vol. 1, pp. 373–377.

[102] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, pp. 239–241, June 2002.

[103] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[104] G. Jongren and M. Skoglund, "Utilizing quantized feedback information in orthogonal space-time block coding," in *Proc. GLOBECOM'00*, Nov. 2000, vol. 2, pp. 995–999.

[105] B. Ozdemir and O. Gurbuz, "A feasible high capacity wireless network architecture with MIMO precoding," *2005 International Conference on Wireless Networks, Communications and Mobile Computing*, Maui, Hawaii, June 2005, vol. 1, pp. 733–738.

[106] L. Sanguinetti and M. Morelli, "Non-linear precoding for MIMO multi-user downlink transmissions with different QoS requirements," in *Proc. EUSIPCO '06*, Florence, Italy, Sep. 2006.

Appendix A Optimal ZF Nonlinear Pre-Processing under the Relative SNR Requirement

It is clear that the optimal processing matrices that minimize α achieve minimum BER for each mobile for the required relative SNR. From (2.32) and the SNR requirement of MS_k, $\alpha = |(\mathbf{H})_k \mathbf{N}_k \mathbf{t}_k|^2 / (\sigma_n^2 \lambda_k)$. Therefore, the solution of $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_M$ should be the one that minimizes $\sum_{k=1}^{M} [|(\mathbf{H})_k \mathbf{N}_k \mathbf{t}_k|^2 / (\sigma_n^2 \lambda_k)]$ under the constraint (2.29). Using the method of Lagrange multipliers, it can be shown that the optimal solution for the processing matrices should be in the form of (2.35)–(2.37) while using the power allocation factors defined in (2.39).

It is noted that the problem of designing processing matrices optimal under different relative SNR requirements has been considered in [106]. The approaches used in this thesis and in [106] are different, but yield the same result.

Appendix B MMSE Nonlinear Pre-Processing Algorithm

First, we consider the case when $M = N$. In this case, $\mathbf{H}^{\dagger} = \mathbf{H}^{-1}$. Also, due to our assumption $\mathbf{F}^H\mathbf{F} = \mathbf{I}$, $\mathbf{F}$ should be a unitary matrix, i.e., $\mathbf{F}^{-1} = \mathbf{F}^H$. Therefore, (2.49) becomes

$$\mathbf{G}(\mathbf{H}\mathbf{H}^H + \xi\mathbf{I}) = \mathbf{B}\mathbf{F}^H\mathbf{H}^H. \tag{B.1}$$

As a result,

$$\mathbf{F}^H = \mathbf{B}^{-1}\mathbf{G}(\mathbf{H} + \xi\mathbf{H}^{-H}). \tag{B.2}$$

Through $\mathbf{F}^H\mathbf{F} = \mathbf{I}$, we get

$$\mathbf{G}^{-1}\mathbf{B}\mathbf{B}^H\mathbf{G}^{-H} = (\mathbf{H} + \xi\mathbf{H}^{-H})(\mathbf{H}^H + \xi\mathbf{H}^{-1}). \tag{B.3}$$

Let $\mathbf{R} = \mathbf{G}^{-1}\mathbf{B}$,

$$\mathbf{R}\mathbf{R}^H = (\mathbf{H} + \xi\mathbf{H}^{-H})(\mathbf{H}^H + \xi\mathbf{H}^{-1}). \tag{B.4}$$

Thus, $\mathbf{R}$ can be obtained from the Cholesky factorization of $(\mathbf{H} + \xi\mathbf{H}^{-H})(\mathbf{H}^H + \xi\mathbf{H}^{-1})$ and is an $M \times M$ lower triangular matrix. $\mathbf{G}$ and $\mathbf{B}$ can be found by

$$\mathbf{G}=\text{diag}[[\mathbf{R}]_{11}^{-1}, [\mathbf{R}]_{22}^{-1}, ..., [\mathbf{R}]_{MM}^{-1}] \tag{B.5}$$

$$\mathbf{B} = \mathbf{G}\mathbf{R}. \tag{B.6}$$

Then, from (B.2), we get

$$\mathbf{F} = (\mathbf{H}^H + \xi\mathbf{H}^{-1})\mathbf{R}^{-H}. \tag{B.7}$$

It can be seen that the above result of $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{G}$ can also be obtained through the QR factorization of $\mathbf{H} + \xi\mathbf{H}^{-H}$. If

$$\mathbf{H} + \xi\mathbf{H}^{-H} = \mathbf{Q}'\mathbf{R}' \tag{B.8}$$

where $\mathbf{R}'$ is an $M \times M$ upper triangular matrix,

$$\mathbf{G}=\mathrm{diag}[[\mathbf{R}']_{11}^{-1},[\mathbf{R}']_{22}^{-1},\ldots,[\mathbf{R}']_{MM}^{-1}] \tag{B.9}$$

$$\mathbf{B}=\mathbf{G}\mathbf{R}'^{H} \tag{B.10}$$

$$\mathbf{F}=\mathbf{Q}'. \tag{B.11}$$

Without loss of generality, assuming the matrix **G** has positive diagonal elements, it is clear that matrices **B**, **F**, and **G** derived by the above equations are unique. Also, it can be seen that the result in (B.4)–(B.11) is just the special case of (2.50)–(2.57) when $M = N$. Therefore, when $M = N$, the **B**, **F**, and **G** found by (2.50)–(2.53) or (2.54)–(2.57) are the unique solution that satisfies (2.49).

Next, let us consider the case when $M \neq N$. Since we have assumed that $N \geq M$, in this case $N > M$. Therefore, $\mathbf{H}^{\dagger} = \mathbf{H}^{H}(\mathbf{H}\mathbf{H}^{H})^{-1}$. Since the results from (2.50)–(2.53) and (2.54)–(2.57) are equivalent, in the following, we only prove that the results from (2.54)–(2.57) satisfy (2.49).

From (2.54)–(2.57), we get

$$\begin{aligned}\mathbf{H}^{H}+\xi\mathbf{H}^{\dagger} &= \mathbf{Q}'\mathbf{R}' \\ &= \mathbf{F}\mathbf{R}' \\ &= \mathbf{F}\mathbf{B}^{H}\mathbf{G}^{-H}\end{aligned} \tag{B.12}$$

so

$$\mathbf{F}=(\mathbf{H}^{H}+\xi\mathbf{H}^{\dagger})\mathbf{R}^{-1}=\mathbf{H}^{H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{R}^{-1} \tag{B.13}$$

and

$$\begin{aligned}\mathbf{F}^{H} &= \mathbf{R}^{-H}(\mathbf{H}+\xi\mathbf{H}^{\dagger H}) \\ &= \mathbf{R}^{-H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{H} \\ &= \mathbf{B}^{-1}\mathbf{G}(\mathbf{H}+\xi\mathbf{H}^{\dagger H})\end{aligned} \tag{B.14}$$

Using the $\mathbf{F}^{H}$ given by (B.14), the r.h.s. of (2.49) is

$$\mathbf{B}\mathbf{F}^{H}\mathbf{H}^{H}=\mathbf{G}(\mathbf{H}\mathbf{H}^{H}+\xi\mathbf{I}). \tag{B.15}$$

Since we have assumed that **F** has orthonormal columns,

$$\mathbf{F}^{H}\mathbf{F}=\mathbf{I}. \tag{B.16}$$

Using the **F** and $\mathbf{F}^{H}$ given by (B.13) and (B.14), it can be found that

$$\mathbf{R}^{-H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{H}\mathbf{H}^{H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{R}^{-1}=\mathbf{I}, \tag{B.17}$$

so

$$\mathbf{R}'^{H}\mathbf{R}' = (\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{H}\mathbf{H}^{H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1}), \quad \text{(B.18)}$$

$$\mathbf{R}^{-1}\mathbf{R}^{-H} = (\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})^{-1}(\mathbf{H}\mathbf{H}^{H})^{-1}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})^{-1}. \quad \text{(B.19)}$$

Also, using $\mathbf{F}$ and $\mathbf{F}^{H}$ given by (B.13) and (B.14), $\mathbf{H}\mathbf{F}\mathbf{F}^{H}\mathbf{H}^{H}$ in the l.h.s. of (2.49) becomes

$$\mathbf{H}\mathbf{F}\mathbf{F}^{H}\mathbf{H}^{H} = \mathbf{H}\mathbf{H}^{H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{R}^{-1}\mathbf{R}^{-H}(\mathbf{I}+\xi(\mathbf{H}\mathbf{H}^{H})^{-1})\mathbf{H}\mathbf{H}^{H}. \quad \text{(B.20)}$$

Using the result for $\mathbf{R}^{-1}\mathbf{R}^{-H}$ given by (B.19), (B.20) becomes

$$\mathbf{H}\mathbf{F}\mathbf{F}^{H}\mathbf{H}^{H} = \mathbf{H}\mathbf{H}^{H}. \quad \text{(B.21)}$$

Therefore, the l.h.s. of (2.49) is

$$\mathbf{G}(\mathbf{H}\mathbf{F}\mathbf{F}^{H}\mathbf{H}^{H}+\xi\mathbf{I}) = \mathbf{G}(\mathbf{H}\mathbf{H}^{H}+\xi\mathbf{I}). \quad \text{(B.22)}$$

From (B.22) and (B.15), we see that the l.h.s. of (2.49) equals the r.h.s. of (2.49). Therefore, (2.49) holds. We can draw the conclusion that processing matrices **B**, **F**, and **G** given by (2.50)–(2.53) or (2.54)–(2.57), satisfy (2.49).

Appendix C Derivation of the Error Covariance Matrix for the MMSE Nonlinear Pre-Processing Algorithm

The error covariance matrix for the nonlinear pre-processing structure shown in Figure 2.2 is in the form of (2.45). With **B**, **F**, and **G** given by (2.50)–(2.53) or (2.54)–(2.57), we get

$$\mathbf{GHF}-\mathbf{B}=\mathbf{GHH}^{H}(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})\mathbf{R}'^{-1}-\mathbf{GR}'^{H}. \tag{C.1}$$

Here, the expression of **F** in (B.13) is used. Then, from (B.19), we get

$$\mathbf{R}'^{-1}=(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})^{-1}(\mathbf{HH}^{H})^{-1}(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})^{-1}\mathbf{R}'^{H}, \tag{C.2}$$

and

$$\begin{aligned}\mathbf{GHF}-\mathbf{B}&=\mathbf{G}(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})^{-1}\mathbf{R}'^{H}-\mathbf{GR}'^{H}\\&=-\xi\mathbf{G}(\mathbf{HH}^{H})^{-1}(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})^{-1}\mathbf{R}'^{H}\end{aligned} \tag{C.3}$$

From (2.45)

$$\begin{aligned}&\mathbf{\Phi}_{ee}=\sigma_{\tilde{a}}^{2}(\mathbf{GHF}-\mathbf{B})(\mathbf{GHF}-\mathbf{B})^{H}+\sigma_{n}^{2}\mathbf{GG}\\&=\sigma_{n}^{2}\xi\mathbf{G}(\mathbf{HH}^{H})^{-1}(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})^{-1}\mathbf{R}'^{H}\,\mathbf{R}'(\mathbf{I}+\xi(\mathbf{HH}^{H})^{-1})^{-1}(\mathbf{HH}^{H})^{-1}\mathbf{G}\\&+\sigma_{n}^{2}\mathbf{GG}\end{aligned} \tag{C.4}$$

Using the result of $\mathbf{R}'^{H}\mathbf{R}'$ given by (B.18),

$$\mathbf{\Phi}_{ee}=\sigma_{n}^{2}\xi\mathbf{G}(\mathbf{HH}^{H})^{-1}\mathbf{G}+\sigma_{n}^{2}\mathbf{G}^{2}. \tag{C.5}$$

Appendix D Proof of the Optimal Ordering Lemma

First, we assume the order generated by the "best-first" ordering method is $L \equiv \{l_1, l_2, ..., l_M\}$. Lemma D.1, which is straightforward from the concept of "best-worst" order, is needed in the proof.

Lemma D.1: Assume the worst performance parameter when order L used is p_{l_r}. Order L is the "best-worst" order, if: for any other order $Q \equiv \{q_1, q_2, ..., q_M\}$, there is an performance parameter p_{q_i} that is the same as or worse than p_{l_r}.

For any element l_i (i = 1,...,M) of order L, we define a *Before Set* $\hat{B}_{l_i} = \{l_1, l_2, ..., l_{i-1}\}$ that includes all the elements located before l_i in L and an *After Set* $\hat{A}_{l_i} = \{l_{i+1}, l_{i+2}, ..., l_M\}$ that includes all the elements located after l_i in L. Note that $\hat{B}_{l_i}$ or $\hat{A}_{l_i}$ can be null set if $i = 1$ or $i = M$.

Then, find the element in L that is associated with the worst performance parameter. We assume this element is l_r and the worst performance parameter is p_{l_r}.

Case 1: Let's consider any other order Q, in which l_r is still in its rth position ($q_r = l_r$) and the Before Set $\hat{B}_{q_r}$ and After Set $\hat{A}_{q_r}$ are permutations of $\hat{B}_{l_r}$ and $\hat{A}_{l_r}$ respectively, i.e.,

$$Q \equiv \{\text{perm}(\hat{B}_{l_r}), l_r, \text{perm}(\hat{A}_{l_r})\} = \{\text{perm}(l_1, l_2, ..., l_{r-1}), l_r, \text{perm}(l_{r+1}, l_{r+2}, ..., l_M)\}$$

where perm($\hat{S}$) represents one permutation of the set $\hat{S}$.

Through Condition 1, it is obvious that $p_{q_r} = p_{l_r}$.

Case 2: Let's consider any other order Q, in which l_r is still located before all the elements of its After Set $\hat{A}_{l_r}$, but one or some elements from the Before Set $\hat{B}_{l_r}$ are moved after l_r.

$$Q \equiv \{q_1, q_2, ..., q_M\} = \{l_1, ..., l_{r-2}, l_r, l_{r-1}, l_{r+1}, ...l_M\}$$

gives a simple example of this case.

If $q_d = l_r$ in this new order Q, then $d < r$ and the constraint set $\hat{C}_{q_d} \supseteq \hat{C}_{l_r}$. Apparently, from Condition 2, p_{q_d} must be worse than or the same as p_{l_r}.

Case 3: Let's consider any other order Q, in which the elements of the constraint set $\hat{C}_{l_r}$ of l_r are distributed in Q (these elements are not needed to be located next to each other) and l_r is not located before all the other elements of $\hat{C}_{l_r}$. Two simple examples of this case are:

$$Q \equiv \{q_1, q_2, ..., q_M\} = \{l_1, ..., l_{r-1}, l_{r+1}, l_r, l_{r+2}, ..., l_M\}$$

$$Q \equiv \{q_1, q_2, ..., q_M\} = \{l_1, ..., l_{r-2}, l_{r+1}, l_r, l_{r-1}, l_{r+2}, ..., l_M\}$$

Assuming l_k $(k > r)$ is the first element of $\hat{C}_{l_r}$ in Q, we assume it is the dth element of Q, i.e., $q_d = l_k$ $(d \leqslant r)$. First, we form a new order

$$V \equiv \{..., v_r, v_{r+1}, ...v_M\} = \{..., l_k, l_r, ..., l_{k-1}, l_{k+1}, ..., l_M\},$$

where $v_r = l_k = q_d$ and $\{v_{r+1}, ...v_M\} = \{l_r, ..., l_{k-1}, l_{k+1}, ..., l_M\}$. It is apparent that $\hat{C}_{q_d} \supseteq \hat{C}_{v_r}$. Since $v_r = q_d$ and through the Condition 2, p_{q_d} must be worse than or the same as p_{v_r}. Next, we see that the constraint set $\hat{C}_{l_r} = \hat{C}_{v_r}$. Since the "best-first" ordering ensures that l_r is the best one for the constraint set $\hat{C}_{l_r}$, p_{v_r} must be worse than or the same as p_{l_r}. As a result, p_{q_d} must be worse than or the same as p_{l_r}.

Since all the other orders can be covered by cases 1, 2, and 3, from Lemma D.1 we can draw the conclusion that order L is optimal in the "best-worst" sense.

Appendix E Reduced Complexity Method Used in the Ordering of the Nonlinear Pre-Processing Algorithms

When $l_i = u_j$ (j = 1,…,P), l_{i+1} ,…, l_M should be a permutation of $u_1,\ldots,u_{j-1},u_{j+1},\ldots,u_P$. Through Condition 1 of the optimal ordering lemma, we know different order of $u_1,\ldots,u_{j-1},u_{j+1},\ldots,u_P$ does not influence the value of $E[|\tilde{n}_{i,(u_j)}|^2]$ or $\varepsilon_{i,(u_j)}$, so we only consider the order $\{ l_i , l_{i+1} ,\ldots, l_M \} = \{u_j,u_P,\ldots,u_{j+1},u_{j-1},\ldots,u_1\}$.

For the conventional ZF nonlinear pre-processing algorithm, we create a matrix

$$\mathbf{H}_{U(u_j)} = [(\mathbf{H})^T_{u_1},\ldots,(\mathbf{H})^T_{u_{j-1}},(\mathbf{H})^T_{u_{j+1}},\ldots,(\mathbf{H})^T_{u_P},(\tilde{\mathbf{H}})^T_{u_j}]^T \tag{E.1}$$

and do the Cholesky factorization with respect to $\mathbf{H}_{U(u_j)}\mathbf{H}^H_{U(u_j)}$, i.e.,

$$\mathbf{H}_{U(u_j)}\mathbf{H}^H_{U(u_j)} = \mathbf{R}_{U(u_j)}\mathbf{R}^H_{U(u_j)}. \tag{E.2}$$

Hence

$$(\mathbf{H}_{U(u_j)}\mathbf{H}^H_{U(u_j)})^{-1} = \mathbf{R}^{-H}_{U(u_j)}\mathbf{R}^{-1}_{U(u_j)}. \tag{E.3}$$

Here, $\mathbf{R}_{U(u_j)}$ is a lower triangular matrix with diagonal elements $[\mathbf{R}_{U(u_j)}]_{11},[\hat{\mathbf{R}}_{U(u_j)}]_{22},\ldots,[\mathbf{R}_{U(u_j)}]_{PP}$, so $\mathbf{R}^{-1}_{U(u_j)}$ is also a lower triangular matrix with diagonal elements $[\mathbf{R}_{U(u_j)}]^{-1}_{11},[\mathbf{R}_{U(u_j)}]^{-1}_{22},\ldots,[\mathbf{R}_{U(u_j)}]^{-1}_{PP}$. From (2.15)–(2.18) and (2.24), we know

$$E[|\tilde{n}_{i,(u_j)}|^2] = \sigma_n^2[\mathbf{R}_{U(u_j)}]^2_{PP} \tag{E.4}$$

By (E.3) and (E.4) and using the feature that $\mathbf{R}_{U(u_j)}^{-1}$ is a lower triangular matrix, it can be found that

$$\begin{aligned} \mathrm{E}[|\tilde{n}_{l,(u_j)}|^2] &= \sigma_n^2[(\mathbf{H}_{U(u_j)}\mathbf{H}_{U(u_j)}^H)^{-1}]_{PP} \\ &= \sigma_n^2[(\mathbf{H}_U\mathbf{H}_U^H)^{-1}]_{jj} \end{aligned} \tag{E.5}$$

where $\mathbf{H}_U$ is defined by (3.7). Thus,

$$(\mathrm{E}[|\tilde{n}_{l,(u_1)}|^2], \mathrm{E}[|\tilde{n}_{l,(u_2)}|^2], \ldots, \mathrm{E}[|\tilde{n}_{l,(u_P)}|^2]) = \sigma_n^2 \mathrm{diag}((\mathbf{H}_U\mathbf{H}_U^H)^{-1}) \tag{E.6}$$

holds as in (3.9).

For the MMSE nonlinear pre-processing algorithm, we create $\mathbf{H}_{U(u_j)}$ as in (E.1), $\mathbf{D}_{U(u_j)}$ as

$$\mathbf{D}_{U(u_j)} = [[\mathbf{H}^\dagger]_{u_1}, \ldots, [\mathbf{H}^\dagger]_{u_{j-1}}, [\mathbf{H}^\dagger]_{u_{j+1}}, \ldots, [\mathbf{H}^\dagger]_{u_P}, [\mathbf{H}^\dagger]_{u_j}], \tag{E.7}$$

and $\mathbf{S}$ as

$$\mathbf{S} = (\mathbf{H}\mathbf{H}^H)^{-1}. \tag{E.8}$$

Then, we generate $\mathbf{G}_{U(u_j)}$, $\mathbf{R}_{U(u_j)}$, and $\mathbf{B}_{U(u_j)}$ through the Cholesky factorization, i.e.,

$$\begin{aligned} &(\mathbf{H}_{U(u_j)} + \xi\mathbf{D}_{U(u_j)}^H)(\mathbf{H}_{U(u_j)}^H + \xi\mathbf{D}_{U(u_j)}) \\ &= \mathbf{R}_{U(u_j)}\mathbf{R}_{U(u_j)}^H \\ &= \mathbf{G}_{U(u_j)}^{-1}\mathbf{B}_{U(u_j)}\mathbf{B}_{U(u_j)}^H\mathbf{G}_{U(u_j)}^{-1} \end{aligned} \tag{E.9}$$

Here, $\mathbf{G}_{U(u_j)}$ is a diagonal matrix that makes $\mathbf{B}_{U(u_j)}$ a unit lower triangular matrix, i.e.,

$$\begin{aligned} \mathbf{G}_{U(u_j)} &= \mathrm{diag}[[\mathbf{G}_{U(u_j)}]_{11}, [\mathbf{G}_{U(u_j)}]_{22}, \ldots, [\mathbf{G}_{U(u_j)}]_{PP}] \\ &= \mathrm{diag}[[\mathbf{R}_{U(u_j)}]_{11}^{-1}, [\mathbf{R}_{U(u_j)}]_{22}^{-1}, \ldots, [\mathbf{R}_{U(u_j)}]_{PP}^{-1}] \end{aligned} \tag{E.10}$$

where $[\mathbf{R}_{U(u_j)}]_{11}, [\mathbf{R}_{U(u_j)}]_{22}, \ldots, [\mathbf{R}_{U(u_j)}]_{PP}$ are the diagonal elements of $\mathbf{R}_{U(u_j)}$.

Through (2.54)–(2.58), it can be found that the error variance $\varepsilon_{l,(u_j)}$ is

$$\varepsilon_{l,(u_j)} = \sigma_n^2[\mathbf{G}_{U(u_j)}]_{PP}^2 + \sigma_n^2\xi[\mathbf{G}_{U(u_j)}]_{PP}^2[\mathbf{S}]_{u_j u_j}. \tag{E.11}$$

Assuming

$$\mathbf{T}_{u_j} = [(\mathbf{H}_{U(u_j)} + \xi\mathbf{D}_{U(u_j)}^H)(\mathbf{H}_{U(u_j)}^H + \xi\mathbf{D}_{U(u_j)})]^{-1}, \tag{E.12}$$

through (E.9)

$$\mathbf{T}_{u_j} = \mathbf{G}_{U(u_j)}\mathbf{B}_{U(u_j)}^{-H}\mathbf{B}_{U(u_j)}^{-1}\mathbf{G}_{U(u_j)} \tag{E.13}$$

$$[\mathbf{G}_{U(u_j)}]_{PP}^2 = [\mathbf{T}_{u_j}]_{PP}. \tag{E.14}$$

Here the feature that $\mathbf{B}_{U(u_j)}$ is a unit lower triangular matrix has been used. Therefore, (E.11) becomes

$$\varepsilon_{l_i(u_j)} = \sigma_n^2[\mathbf{T}_{u_j}]_{PP} + \sigma_n^2\xi[\mathbf{T}_{u_j}]_{PP}[\mathbf{S}]_{u_j u_j} \tag{E.15}$$

Also, let

$$\mathbf{T} = [(\mathbf{H}_U + \xi\mathbf{D}_U^H)(\mathbf{H}_U^H + \xi\mathbf{D}_U)]^{-1} \tag{E.16}$$

where $\mathbf{H}_U$ and $\mathbf{D}_U$ are defined in (3.7) and (3.8). It can be found that

$$[\mathbf{T}_{u_j}]_{PP} = [\mathbf{T}]_{jj} \tag{E.17}$$

As a result,

$$\varepsilon_{l_i(u_j)} = \sigma_n^2[\mathbf{T}]_{jj} + \sigma_n^2\xi[\mathbf{T}]_{jj}[\mathbf{S}]_{u_j u_j}. \tag{E.18}$$

Hence the result in (3.10)–(3.12) is given.

Appendix F Proof that the ZF Nonlinear Joint Tx-Rx Processing Algorithm Achieves the Multi-user MIMO Decomposition

First, we represent the matrix $\mathbf{B}$ in (4.34) as $\mathbf{B} = \mathbf{B}_D\mathbf{B}_U$ and the product $\mathbf{HF}$ in (4.41) as $\mathbf{HF} = (\mathbf{HF})_D(\mathbf{HF})_U$, where $\mathbf{B}_D$ and $(\mathbf{HF})_D$ are block diagonal matrices defined as

$$\mathbf{B}_D \equiv \begin{bmatrix} \mathbf{B}_{1,1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{K,K} \end{bmatrix}, \tag{F.1}$$

$$(\mathbf{HF})_D \equiv \begin{bmatrix} \mathbf{H}_1\mathbf{F}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2\mathbf{F}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}_K\mathbf{F}_K \end{bmatrix}. \tag{F.2}$$

From (4.38), one can see that the ZF criterion requires $\mathbf{DHF} = \mathbf{B}$. Since $\mathbf{D}$ is a block diagonal matrix defined in (4.35), it is clear that $\mathbf{D}(\mathbf{HF})_D = \mathbf{B}_D$ and $(\mathbf{HF})_U = \mathbf{B}_U$. The received signal vector $\mathbf{y}$ is then found to be

$$\mathbf{y} = \mathbf{HFB}^{-1}\mathbf{v} + \mathbf{n} = \mathbf{HFB}_U^{-1}\mathbf{B}_D^{-1}\mathbf{v} + \mathbf{n} \tag{F.3}$$

where $\mathbf{B} = \mathbf{B}_D\mathbf{B}_U$ is used. Since $(\mathbf{HF})_U = \mathbf{B}_U$,

$$\mathbf{y} = \mathbf{HF}(\mathbf{HF})_U^{-1}\mathbf{B}_D^{-1}\mathbf{v} + \mathbf{n}. \tag{F.4}$$

Finally, using $\mathbf{HF} = (\mathbf{HF})_D(\mathbf{HF})_U$,

$$\mathbf{y} = (\mathbf{HF})_D\mathbf{B}_D^{-1}\mathbf{v} + \mathbf{n}. \tag{F.5}$$

Therefore, (4.45) and (4.46) can be obtained using the definitions of $\mathbf{B}_D$ and $(\mathbf{HF})_D$ in (F.1) and (F.2), so Lemma 4.1 is proved.

Appendix G ZF Nonlinear Joint Tx-Rx Processing Algorithm

The following lemma from [90] and [94] is used in finding the solution of (4.50).

Lemma G.1: Given a rank K_r matrix $\mathbf{X}$ whose size is $M \times N$, the solution of

$$\begin{aligned} &\max_{\mathbf{F},\mathbf{G}} \min\{|[\mathbf{U}]_{ii}| : 1 \le i \le K_r\} \\ &\text{subject to} \quad \mathbf{U} = \mathbf{G}^H \mathbf{X} \mathbf{F}, \\ &\qquad\qquad \mathbf{U} \text{ is an upper triangular matrix,} \\ &\qquad\qquad \mathrm{tr}\{\mathbf{G}^H \mathbf{G}\} \le K_r, \\ &\qquad\qquad \mathrm{tr}\{\mathbf{F}^H \mathbf{F}\} \le K_r, \end{aligned} \tag{G.1}$$

is given by $\mathbf{G} = \mathbf{Q}$, $\mathbf{U} = \mathbf{R}$, and $\mathbf{F} = \mathbf{P}$, where $\mathbf{QRP}^H$ is the GMD of $\mathbf{X}$.

To prove that (4.51) achieves the solution of (4.50), first, let us introduce matrices $\mathbf{T}_k$ and $\mathbf{D}_k'$ as

$$\mathbf{T}_k = \mathrm{diag}[\sqrt{\sigma_n^2 / \sigma_{\tilde{n}_k^{(1)}}^2}, \sqrt{\sigma_n^2 / \sigma_{\tilde{n}_k^{(2)}}^2}, \ldots, \sqrt{\sigma_n^2 / \sigma_{\tilde{n}_k^{(m_k)}}^2}], \tag{G.2}$$

and

$$\mathbf{D}_k' \equiv \mathbf{T}_k \mathbf{D}_k. \tag{G.3}$$

From (4.49), (G.2), and (G.3), it can be found that

$$\mathrm{diag}(\mathbf{D}_k' \mathbf{D}_k'^H) = (1,1,\ldots,1)^T, \tag{G.4}$$

so

$$\mathrm{tr}(\mathbf{D}_k' \mathbf{D}_k'^H) = m_k. \tag{G.5}$$

Using (G.3), one can see that (4.47) can be written as

$$\mathbf{D}_k' \mathbf{H}_k \mathbf{N}_k \mathbf{A}_k = \mathbf{T}_k \mathbf{B}_{k,k}, \tag{G.6}$$

where $\mathbf{T}_k\mathbf{B}_{k,k}$ is an upper triangular matrix and we use $\mathbf{U}_k$ to represent it. The diagonal elements of $\mathbf{U}_k$ should be $\mathrm{diag}(\mathbf{T}_k)$.

Clearly, (4.50) can be written equivalently as

$$\max_{\mathbf{A}_k,\mathbf{D}_k'} \min\{|[\mathbf{U}_k]_{ii}| : 1 \le i \le m_k\}$$

$$\text{subject to} \quad \mathbf{D}_k'\mathbf{H}_k\mathbf{N}_k\mathbf{A}_k = \mathbf{U}_k, \tag{G.7}$$

$$\mathbf{U}_k \text{ is an upper triangular matrix,}$$

$$\mathrm{tr}\{\mathbf{A}_k\mathbf{A}_k^H\} = m_k,$$

$$\mathrm{tr}(\mathbf{D}_k'\mathbf{D}_k'^H) = m_k.$$

Finally, from Lemma G.1, it can be seen that the solution of (G.7) is given by (4.51). Therefore, the solution of (4.50) is given by (4.51).

Appendix H Optimal Order in the Maximal Diversity Order Sense for the ZF Nonlinear Joint Tx-Rx Processing Algorithm

If the $\overline{L}$ satisfying (4.65) is used, we assume mobile $\overline{l}_r$ is the mobile with the minimum diversity order, so the diversity order of the system is

$$d^{(\overline{L})} = d_{MS_{\overline{l}_r}}. \tag{H.1}$$

It can be easily found that $m_{\overline{l}_k} > m_{\overline{l}_r}$ for any $k < r$.

For any other order $Q \equiv \{q_1, q_2, \ldots, q_K\}$, let us consider the diversity order of the mobile q_r. Since order $\overline{L}$ satisfies (4.65), one can see that

$$(N - \sum_{i=r}^{K} m_{q_i} + 1) \le (N - \sum_{i=r}^{K} m_{\overline{l}_i} + 1) \tag{H.2}$$

must hold.

If $m_{q_r} \le m_{\overline{l}_r}$, from (4.63) and (H.2), it is clear that

$$d_{MS_{q_r}} \le d_{MS_{\overline{l}_r}}. \tag{H.3}$$

If $m_{q_r} > m_{\overline{l}_r}$, it can be found that there must be a mobile q_t whose number of antennas is smaller than or equal to that of mobile $\overline{l}_r$, i.e.,

$$m_{q_t} \le m_{\overline{l}_r}, \tag{H.4}$$

and with $t < r$. In this case, since $t < r$

$$(N - \sum_{i=t}^{K} m_{q_i} + 1) < (N - \sum_{i=r}^{K} m_{q_i} + 1) \le (N - \sum_{i=r}^{K} m_{\overline{l}_i} + 1). \tag{H.5}$$

From (4.63), (H.4), and (H.5), it is clear that

$$d_{MS_{q_t}} < d_{MS_{\overline{l}_r}}. \tag{H.6}$$

Therefore, when any other order Q is used, there is a mobile whose diversity order is smaller than or the same as the system diversity order when order $\overline{L}$ is used. Since the system diversity order is the minimum diversity order of all the mobiles, the system diversity order when any other order Q is used must be smaller than or the same as the system diversity order when order $\overline{L}$ is used. As a conclusion the order $\overline{L}$ is optimal in the maximal diversity order sense.

Appendix I Application of the "Best-First" Ordering for the ZF Nonlinear Joint Tx-Rx Processing Algorithm

In this part, we will prove that the "best-first" ordering can achieve the optimal order in the minimax noise variance sense for the group of mobiles with equal number of antennas when the ZF nonlinear joint Tx-Rx processing algorithm is used.

In Section 3.3, an *optimal ordering lemma* is introduced. According to this lemma, if the following two conditions are satisfied, the "best-first" ordering can achieve the optimal order in the minimax noise variance sense for $\overline{\mathbf{l}}_i \equiv \{\overline{l}_i, \overline{l}_{i+1}, ..., \overline{l}_{i+p}\}$.

Condition 1: For any two orders $A \equiv \{A_1, A_2, ..., A_p\}$ and $B \equiv \{B_1, B_2, ..., B_p\}$ of $\overline{\mathbf{l}}_i \equiv \{\overline{l}_i, \overline{l}_{i+1}, ..., \overline{l}_{i+p}\}$, if $A_t = B_t$ and the *constraint sets* $\hat{C}_{A_t}$ and $\hat{C}_{B_t}$ are made up of the same elements, then $\sigma^2_{\overline{n}_{A_t}} = \sigma^2_{\overline{n}_{B_t}}$.

Condition 2: For any two orders A and B of $\overline{\mathbf{l}}_i$, if $A_t = B_j$ and the constraint set $\hat{C}_{A_t} \subseteq \hat{C}_{B_j}$, then $\sigma^2_{\overline{n}_{A_t}} \le \sigma^2_{\overline{n}_{B_j}}$.

The definition of the constraint set can be found in Section 3.3.

Since the mobiles $\overline{l}_i, \overline{l}_{i+1}, ..., \overline{l}_{i+p}$ have the same number of antennas, we use b to represent this number. From (4.55), one can find that

$$\sigma^2_{\overline{n}_{A_t}} = \sigma_n^2 \overline{\sigma}^{-2}_{A_t} = \sigma_n^2 \left(\prod_{j=1}^{b} \sigma^{(j)}_{A_t}\right)^{-2/b}, \tag{I.1}$$

where $\sigma^{(j)}_{A_t}$ ($j = 1, ..., b$) are the singular values of $\mathbf{H}_{A_t}\mathbf{N}_{A_t}$, so

$$\prod_{j=1}^{b} \sigma_{A_t}^{(j)2} = \det(\mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H). \tag{I.2}$$

$\mathbf{N}_{A_t}$ can be found from the SVD of $\bar{\mathbf{H}}_{A_t}$, i.e.,

$$\bar{\mathbf{H}}_{A_t} = \tilde{\mathbf{U}}_{A_t}\begin{bmatrix}\boldsymbol{\Sigma}_{A_t} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{N}}_{A_t}^H \\ \mathbf{N}_{A_t}^H\end{bmatrix}, \tag{I.3}$$

where $\bar{\mathbf{H}}_{A_t}$ is

$$\bar{\mathbf{H}}_{A_t} \equiv [\mathbf{H}_{A_{t+1}}^T, \ldots, \mathbf{H}_{A_p}^T, \mathbf{H}_{\bar{I}_{t+p+1}}^T, \ldots, \mathbf{H}_{\bar{I}_K}^T]^T. \tag{I.4}$$

For any two orders A and B, if $A_t = B_t$, and $\hat{C}_{A_t}$ and $\hat{C}_{B_t}$ are made up of the same elements, from (I.3) and (I.4), one can see that ran($\mathbf{N}_{A_t}$) = ran($\mathbf{N}_{B_t}$). Also, since $A_t = B_t$, $\mathbf{H}_{A_t} = \mathbf{H}_{B_t}$. Thus,

$$\det(\mathbf{H}_{B_t}\mathbf{N}_{B_t}\mathbf{N}_{B_t}^H\mathbf{H}_{B_t}^H) = \det(\mathbf{H}_{A_t}\mathbf{N}_{B_t}\mathbf{N}_{B_t}^H\mathbf{H}_{A_t}^H) = \det(\mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H). \tag{I.5}$$

From (I.1), (I.2), and (I.5), $\sigma_{\bar{n}_{A_t}}^2 = \sigma_{\bar{n}_{B_t}}^2$. Therefore, the Condition 1 is satisfied.

For any two orders A and B, if $A_t = B_j$, and $\hat{C}_{A_t} \subseteq \hat{C}_{B_j}$, from (I.3) and (I.4) one can see that ran($\mathbf{N}_{B_j}$) is a subspace of ran($\mathbf{N}_{A_t}$). Since $\mathbf{N}_{A_t}^H\mathbf{N}_{A_t} = \mathbf{I}$,

$$\mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H = \mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H. \tag{I.6}$$

Since ran($\mathbf{N}_{B_j}$) is a subspace of ran($\mathbf{N}_{A_t}$), $\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{N}_{B_j} = \mathbf{N}_{B_j}$, so

$$\mathbf{H}_{B_j}\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{H}_{B_j}^H = \mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H \tag{I.7}$$

where $\mathbf{H}_{B_j} = \mathbf{H}_{A_t}$ is used. Using $\mathbf{T}$ to represent $\mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H$, i.e., $\mathbf{T} \equiv \mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H$,

$$\mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H = \mathbf{T}\mathbf{T}^H, \tag{I.8}$$

$$\mathbf{H}_{B_j}\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{H}_{B_j}^H = \mathbf{T}\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{T}^H. \tag{I.9}$$

Assuming the SVD of $\mathbf{T}$ is

$$\mathbf{T} = \mathbf{U}_T\begin{bmatrix}\mathbf{D}_T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_T^H \\ \tilde{\mathbf{V}}_T^H\end{bmatrix}, \tag{I.10}$$

one can have

$$\det(\mathbf{H}_{A_t}\mathbf{N}_{A_t}\mathbf{N}_{A_t}^H\mathbf{H}_{A_t}^H) = \det(\mathbf{T}\mathbf{T}^H) = \det(\mathbf{D}_T^2), \tag{I.11}$$

$$\det(\mathbf{H}_{B_j}\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{H}_{B_j}^H) = \det(\mathbf{T}\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{T}^H) = \det(\mathbf{D}_T^2)\det(\mathbf{V}_T^H\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{V}_T). \quad \text{(I.12)}$$

Using the Hadamard's inequality [85], one get

$$\det(\mathbf{V}_T^H\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{V}_T) \le \prod_{r=1}^{b}((\mathbf{V}_T^H)_r\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H[\mathbf{V}_T]_r). \quad \text{(I.13)}$$

Since for any $r = 1,\ldots,b$,

$$\begin{aligned} & 1-(\mathbf{V}_T^H)_r\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H[\mathbf{V}_T]_r \\ &= (\mathbf{V}_T^H)_r(\mathbf{I}-\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H)[\mathbf{V}_T]_r \\ &= (\mathbf{V}_T^H)_r(\mathbf{I}-\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H)(\mathbf{I}-\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H)[\mathbf{V}_T]_r \\ &\ge 0 \end{aligned} \quad \text{(I.14)}$$

$$(\mathbf{V}_T^H)_r\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H[\mathbf{V}_T]_r \le 1. \quad \text{(I.15)}$$

Therefore,

$$\det(\mathbf{V}_T^H\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{V}_T) \le 1. \quad \text{(I.16)}$$

From (I.11), (I.12), and (I.16)

$$\det(\mathbf{H}_{A_i}\mathbf{N}_{A_i}\mathbf{N}_{A_i}^H\mathbf{H}_{A_i}^H) \ge \det(\mathbf{H}_{B_j}\mathbf{N}_{B_j}\mathbf{N}_{B_j}^H\mathbf{H}_{B_j}^H). \quad \text{(I.17)}$$

From (I.1), (I.2), and (I.17), $\sigma_{\tilde{n}_{A_i}}^2 \le \sigma_{\tilde{n}_{B_j}}^2$. Thus, the Condition 2 is satisfied.

Since the two conditions are satisfied, it is clear that the "best-first" ordering can achieve the optimal order in the minimax noise variance sense for any group of mobiles with equal number of antennas.

Appendix J Design of the Feedback Filter for the MMSE Nonlinear Joint Tx-Rx Processing Algorithm

A similar method as that in deriving the feedback filter of the synchronous MMSE DFD in [63] is used in this proof.

From the definitions of $\mathbf{B}$, $\mathbf{B}_k$, and $\bar{\mathbf{B}}_k$, it can be found that

$$[\bar{\mathbf{B}}_k]_{ij} = 0 \quad i > j, \tag{J.1}$$

$$[\bar{\mathbf{B}}_k]_{ii} = 1 \quad i = 1,\ldots,m_k, \tag{J.2}$$

and $\bar{\mathbf{B}}_k$ is a matrix of size $m_k \times \sum_{i=k}^{K} m_i$.

Since $\bar{\mathbf{R}}_k$ generated from (4.77) is an upper triangular matrix, $\bar{\mathbf{R}}_k^{-1}$ is also an upper triangular matrix with diagonal elements $[\bar{\mathbf{R}}_k^{-1}]_{ii} = [\bar{\mathbf{R}}_k]_{ii}^{-1}$.

We form a matrix

$$\mathbf{U}_k = \bar{\mathbf{B}}_k \bar{\mathbf{R}}_k^{-1}. \tag{J.3}$$

For this matrix, it can be found that

$$[\mathbf{U}_k]_{ij} = 0 \quad i > j, \tag{J.4}$$

$$[\mathbf{U}_k]_{ii} = [\bar{\mathbf{R}}_k^{-1}]_{ii} = [\bar{\mathbf{R}}_k]_{ii}^{-1} \quad i = 1,\ldots,m_k, \tag{J.5}$$

and $\mathbf{U}_k$ is also of size $m_k \times \sum_{i=k}^{K} m_i$

Using the $\mathbf{U}_k$ in (J.3), the $\mathbf{\Phi}_{e_k e_k}$ in (4.75) can be represented as

$$\mathbf{\Phi}_{e_k e_k} = \sigma_n^2 \mathbf{U}_k \mathbf{U}_k^H. \tag{J.6}$$

Therefore,

$$
\begin{aligned}
\mathrm{tr}(\boldsymbol{\Phi}_{\mathbf{e}_k\mathbf{e}_k}) &= \sigma_n^2 \sum_{i=1}^{m_k} \sum_{j=i}^{\sum_{t=k}^{K} m_t} |[\mathbf{U}_k]_{ij}|^2 \\
&\geq \sigma_n^2 \sum_{i=1}^{m_k} |[\mathbf{U}_k]_{ii}|^2 \\
&= \sigma_n^2 \sum_{i=1}^{m_k} |[\bar{\mathbf{R}}_k]_{ii}^{-1}|^2 \\
&= \sigma_n^2 \sum_{i=1}^{m_k} |[\bar{\mathbf{R}}_k']_{ii}^{-1}|^2
\end{aligned}
\tag{J.7}
$$

where we have assumed that $\bar{\mathbf{R}}_k'$ is made up of the 1th,..., m_k th rows of $\bar{\mathbf{R}}_k$.

Clearly, if $\bar{\mathbf{B}}_k$ found by (4.77)–(4.79) is used, $\mathrm{tr}(\boldsymbol{\Phi}_{\mathbf{e}_k\mathbf{e}_k}) = \sigma_n^2 \sum_{i=1}^{m_k} |[\bar{\mathbf{R}}_k']_{ii}^{-1}|^2$.

Therefore, $\bar{\mathbf{B}}_k$ found by (4.77)–(4.79) can minimize $\mathrm{tr}(\boldsymbol{\Phi}_{\mathbf{e}_k\mathbf{e}_k})$.

Appendix K Design of the Feedforward Filter and Derivation of the Noise Variance for the MMSE Nonlinear Joint Tx-Rx Processing Algorithm

Since $\mathbf{A}_k$ is assumed to have orthonormal columns, (4.80) can be written as

$$\mathbf{A}_k^H(\mathbf{N}_k^H\mathbf{H}_k^H\mathbf{H}_k\mathbf{N}_k+\xi\mathbf{I})\mathbf{A}_k=\mathbf{R}_k^H\mathbf{R}_k. \tag{K.1}$$

From (K.1), one can get

$$\prod_{i=1}^{m_k}[\mathbf{R}_k]_{ii}^2=\det(\mathbf{A}_k^H(\mathbf{N}_k^H\mathbf{H}_k^H\mathbf{H}_k\mathbf{N}_k+\xi\mathbf{I})\mathbf{A}_k). \tag{K.2}$$

Using the SVD result of $\mathbf{H}_k\mathbf{N}_k$ in (4.85), we have

$$\mathbf{A}_k^H(\mathbf{N}_k^H\mathbf{H}_k^H\mathbf{H}_k\mathbf{N}_k+\xi\mathbf{I})\mathbf{A}_k=\mathbf{A}_k^H\bar{\mathbf{V}}_k\begin{bmatrix}\boldsymbol{\Sigma}_k^2+\xi\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \xi\mathbf{I}\end{bmatrix}\bar{\mathbf{V}}_k^H\mathbf{A}_k. \tag{K.3}$$

It can be shown that

$$\begin{aligned}\prod_{i=1}^{m_k}[\mathbf{R}_k]_{ii}^2&=\det(\mathbf{A}_k^H\bar{\mathbf{V}}_k\begin{bmatrix}\boldsymbol{\Sigma}_k^2+\xi\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \xi\mathbf{I}\end{bmatrix}\bar{\mathbf{V}}_k^H\mathbf{A}_k)\\ &\le\prod_{i=1}^{m_k}([\mathbf{A}_k]_i^H\bar{\mathbf{V}}_k\begin{bmatrix}\boldsymbol{\Sigma}_k^2+\xi\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \xi\mathbf{I}\end{bmatrix}\bar{\mathbf{V}}_k^H[\mathbf{A}_k]_i)\\ &\le\prod_{i=1}^{m_k}([\boldsymbol{\Sigma}_k]_{ii}^2+\xi)=\prod_{i=1}^{m_k}\sigma_k^{(i)2}\end{aligned} \tag{K.4}$$

where $\sigma_k^{(i)2}$ $(i=1,\ldots,m_k)$ are the singular values of $\mathbf{J}_k$ and $\mathbf{J}_k$ is defined in (4.87).

From (4.80)–(4.82), one can get

$$\mathrm{tr}(\boldsymbol{\Phi}_{\mathbf{e}_k\mathbf{e}_k})=\sigma_n^2\sum_{i=1}^{m_k}[\mathbf{R}_k]_{ii}^{-2}. \tag{K.5}$$

Since (K.4), (K.5), and

$$\sum_{i=1}^{m_k} [\mathbf{R}_k]_{ii}^{-2} \geq m_k \left(\prod_{i=1}^{m_k} [\mathbf{R}_k]_{ii}^{-2}\right)^{1/m_k}, \tag{K.6}$$

one can see that

$$\mathrm{tr}(\boldsymbol{\Phi}_{e_k e_k}) \geq \sigma_n^2 m_k \left(\prod_{i=1}^{m_k} \sigma_k^{(i)2}\right)^{-1/m_k}. \tag{K.7}$$

Therefore,

$$\min\{\mathrm{tr}(\boldsymbol{\Phi}_{e_k e_k})\} = \sigma_n^2 m_k \left(\prod_{i=1}^{m_k} \sigma_k^{(i)2}\right)^{-1/m_k}. \tag{K.8}$$

Finally, from the definition of the GMD in Section 4.3.2, it can be shown that the $\mathbf{A}_k$ found by (4.84)–(4.87) can make $\mathrm{tr}(\boldsymbol{\Phi}_{e_k e_k})$ achieve its minimum value, so it minimizes $\mathrm{tr}(\boldsymbol{\Phi}_{e_k e_k})$. In this case, $\mathbf{R}_k$ will have equal diagonal elements, and one can get

$$[\mathbf{R}_k]_{11}^{-2} = \ldots = [\mathbf{R}_k]_{m_k m_k}^{-2} = \left(\prod_{i=1}^{m_k} \sigma_k^{(i)2}\right)^{-1/m_k}. \tag{K.9}$$

From (4.80)–(4.82) and (K.9), one can see that (4.88) holds.

Appendix L Proof that the Transmitted Power is Allocated Equally in the MMSE Nonlinear Joint Tx-Rx Processing Algorithm

Since $\mathbf{F}_k = \mathbf{N}_k \mathbf{A}_k$,

$$\mathbf{F}_j^H \mathbf{F}_k = \mathbf{A}_j^H \mathbf{N}_j^H \mathbf{N}_k \mathbf{A}_k . \tag{L.1}$$

First, let us consider the case $j = k$. Since the columns of $\mathbf{N}_k$ are the orthonormal basis vectors of null($\bar{\mathbf{H}}_k$), one can see that

$$\mathbf{N}_k^H \mathbf{N}_k = \mathbf{I} . \tag{L.2}$$

Therefore,

$$\mathbf{F}_k^H \mathbf{F}_k = \mathbf{A}_k^H \mathbf{A}_k . \tag{L.3}$$

From (4.83), we know

$$\mathbf{F}_k^H \mathbf{F}_k = \mathbf{I} . \tag{L.4}$$

Next, let us consider the case $j > k$. From (4.84), we know

$$\mathbf{F}_k = \mathbf{N}_k \mathbf{V}_k \mathbf{P}_k , \tag{L.5}$$

so

$$\mathbf{F}_j^H \mathbf{F}_k = \mathbf{P}_j^H \mathbf{V}_j^H \mathbf{N}_j^H \mathbf{N}_k \mathbf{V}_k \mathbf{P}_k . \tag{L.6}$$

Using (4.85) the $\mathbf{V}_j^H$ can be represented as

$$\mathbf{V}_j^H = \boldsymbol{\Sigma}_j^{-1} \mathbf{U}_j^{-1} \mathbf{H}_j \mathbf{N}_j , \tag{L.7}$$

so (L.6) becomes

$$\mathbf{F}_j^H \mathbf{F}_k = \mathbf{P}_j^H \boldsymbol{\Sigma}_j^{-1} \mathbf{U}_j^{-1} \mathbf{H}_j \mathbf{N}_j \mathbf{N}_j^H \mathbf{N}_k \mathbf{V}_k \mathbf{P}_k . \tag{L.8}$$

From the definition of $\bar{\mathbf{H}}_k$ in (4.74), one can see that

$$\text{null}(\bar{\mathbf{H}}_k) \subset \text{null}(\bar{\mathbf{H}}_j) \quad \textit{if} \quad j > k \tag{L.9}$$

Since the columns of $\mathbf{N}_k$ are the orthonormal basis vectors of null($\bar{\mathbf{H}}_k$) and the columns of $\mathbf{N}_j$ are the orthonormal basis vectors of null($\bar{\mathbf{H}}_j$), it can be easily seen that

$$\mathbf{N}_j \mathbf{N}_j^H \mathbf{N}_k = \mathbf{N}_k . \tag{L.10}$$

Therefore, (L.8) becomes

$$\mathbf{F}_j^H \mathbf{F}_k = \mathbf{P}_j^H \boldsymbol{\Sigma}_j^{-1} \mathbf{U}_j^{-1} \mathbf{H}_j \mathbf{N}_k \mathbf{V}_k \mathbf{P}_k . \tag{L.11}$$

Since the columns of $\mathbf{N}_k$ are the orthonormal basis vectors of null($\bar{\mathbf{H}}_k$) and $j > k$, it is clear that

$$\mathbf{H}_j \mathbf{N}_k = \mathbf{0} . \tag{L.12}$$

Therefore,

$$\mathbf{F}_j^H \mathbf{F}_k = \mathbf{0} . \tag{L.13}$$

For the case $j < k$, based on the above proof we know that $\mathbf{F}_k^H \mathbf{F}_j = \mathbf{0}$, so

$$\mathbf{F}_j^H \mathbf{F}_k = \mathbf{0} . \tag{L.14}$$

Finally, from (L.4), (L.13), and (L.14), one can find that

$$\mathbf{F}^H \mathbf{F} = \mathbf{I} . \tag{L.15}$$

As a result, one can see that the transmitted power is allocated equally to all the data streams.